

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ :

H04L 12/56

A3

(11) International Publication Number:

WO 98/27660

(43) International Publication Date:

25 June 1998 (25.06.98)

(21) International Application Number: PCT/US97/23285

(22) International Filing Date: 16 December 1997 (16.12.97)

(30) Priority Data:

08/767,576	16 December 1996 (16.12.96)	US
08/844,171	18 April 1997 (18.04.97)	US
08/901,061	24 July 1997 (24.07.97)	US

(71) Applicant: JUNIPER NETWORKS [US/US]; 3260 Jay Street, Santa Clara, CA 95051 (US).

(72) Inventors: SINDHU, Pradeep, S.; 1557 Montalto Drive, Mountain View, CA 94040 (US). ANAND, Ramalingam, K.; 3096 Toscana Court, San Jose, CA 95135 (US). FERGUSON, Dennis, C.; 203 Orchard Glen Court, Mountain View, CA 94043 (US). LIENCRES, Bjorn, O.; 2731 Greer Road, Palo Alto, CA 94303 (US).

(74) Agents: BOROVOY, Roger, S. et al.; Fish & Richardson P.C., Suite 100, 2200 Sand Hill Road, Menlo Park, CA 94025 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

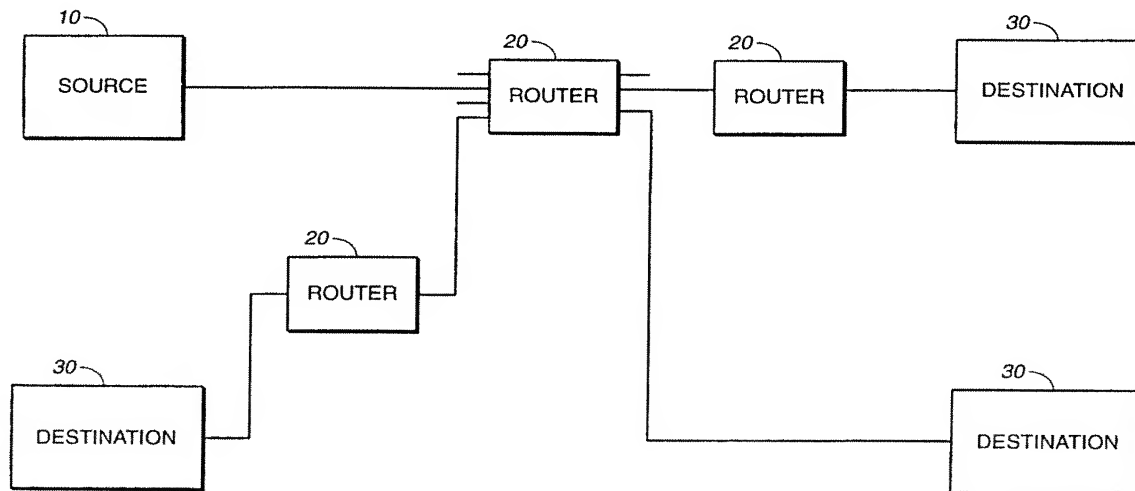
Published

With international search report.

(88) Date of publication of the international search report:

15 October 1998 (15.10.98)

(54) Title: HIGH SPEED SWITCHING DEVICE



(57) Abstract

A router (20) for switching a data packet between a source (10) and destination (30) in a network including a plurality of input ports each including a data handler. The data handler divides a data packet into one or more fixed length cells. The router (20) includes a plurality of output ports at least one of which is for routing the data packet to the destination (30) and a memory divided into a plurality of memory banks. An input switch receives fixed length cells from the input ports and writes a single output switch routes cells received from the memory to an appropriate output port.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

- 1 -

HIGH SPEED SWITCHING DEVICEBackground

5 The present invention relates generally to data routing systems, and more particularly to a method and apparatus for routing packets through a network.

 In packet switch communication systems, a router is a switching device which receives packets containing
10 data or control information on one port, and based on destination information contained within the packet, routes the packet out another port to the destination (or intermediary destination).

 Conventional routers perform this switching
15 function by evaluating header information contained within a first data block in the packet in order to determine the proper output port for a particular packet.

 Efficient switching of packets through the router is of paramount concern. Referring now to Figure 1a, a
20 conventional router includes a plurality of input ports 2 each including an input buffer (memory) 4, a switching device 6 and a plurality of output ports 8.

 Data packets received at an input port 2 are stored at least temporarily, in input buffer 4 while
25 destination information associated with each packet is decoded to determine the appropriate switching through the switching device 6. Obviously, if the decoding process takes too long as compared to the rate at which packets are received, a larger memory element will be
30 required. In addition, the packet may be forced to remain in the input buffer after the destination information is decoded if the switching device cannot make the connection. Blocking refers to a condition in

- 2 -

which a connection cannot be made in the switch due to the unavailability of the desired output port (the port is busy, e.g., routing another packet from a different input port). The size of each input buffer 4 is
5 dependent on a number of factors including: the line input rate, the speed of the look-up process, and the blocking characteristics for the switching device.

Unfortunately, these types of routers are inefficient in a number of respects. Each input port
10 includes a dedicated input buffer and memory sharing between input ports is not provided for in the design. Each input buffer must be sized to meet the maximum throughput requirements for a given port. However, design trades (cost) often necessitate smaller buffers
15 for each port. With the smaller buffers, the possibility arises for packets to be dropped due to blocking conditions. While excess memory capacity typically existed in the router (due to the varied usage of the input ports), no means for taking advantage of the excess
20 is afforded.

To prevent the dropping of packets, designers developed "non-blocking" routers. Referring now to Figure 1b, a conventional "non-blocking" router includes a plurality of input ports 2 each including an input
25 buffer (memory) 4, a switching device 6 and a plurality of output ports 8 each having an output buffer 9. In order to avoid blocking conditions, each output port 8 was configured to include an output buffer 9. Each output port could simultaneously be outputting packets as
30 well as receiving new packets for output at a later time. If the size of the output buffer was sufficiently large, then no data packets would be dropped.

- 3 -

However, these designs are even more inefficient in terms of memory capacity and cost. Again, each output port includes a dedicated output buffer and memory sharing between output ports is not provided for in the design. Each output buffer must be sized to meet the maximum throughput requirements for a given port (in order to maintain its "non-blocking" characteristics). Even more excess memory capacity typically exists in the router (due to the varied usage of the input ports and output ports), yet no means for taking advantage of the excess is afforded. Twice the amount and bandwidth of memory has to be used than required to support the amount of data being moved through these types of devices.

Summary of the Invention

In general, in one aspect, the invention provides a router for switching a data packet between a source and destination in a network including a plurality of input ports each including a data handler. The data handler divides a data packet into one or more fixed length cells. The router includes a plurality of output ports at least one of which is for routing the data packet to the destination and a memory divided into a plurality of memory banks. A input switch receives fixed length cells from the input ports and writes a single cell in a cell slot time span to each memory bank. An output switch routes cells received from the memory to an appropriate output port.

Aspects of the invention include the following features. The input switch includes a linking engine for linking cells in the data packet to allow retrieval of the data packet from non-contiguous locations in the memory.

- 4 -

The router further includes an indirect cell generator for generating one or more indirect cells. The linking engine tracks the location in the memory where consecutive cells of the data packet are stored and
5 provides an address in memory of each cell in the data packet for storage in indirect cells.

The input switch time division multiplexes the writing of data packets to the memory such that consecutive cells from the input port are written to
10 consecutive banks in the memory. The input switch includes a key reading engine for extracting key information from a first cell received at the input switch associated with the data packet. The router further includes a controller coupled to the input switch
15 for receiving the key information therefrom. The controller decodes destination information from the key information received from the input switch and outputs a notification defining a routing of the data packet from the memory to the output port.

20 The output port includes a result processor for receiving the notification from the controller and initiates a transfer of the data packet from the memory to the output port. The input switch includes a reservation table for scheduling transfers from the
25 memory to the output switch. The output switch routes the notification to the output port and thereafter the output port issues a request to the input switch to transfer the data packet from memory to the output port through the output switch. The request from the output
30 port is stored in the reservation table. The requests to transfer cells from memory to the output switch are time domain multiplexed so that during one cell slot time span at most a single read request is issued to each bank in

- 5 -

the memory for servicing. The memory outputs at most a single cell per bank in one cell slot time span.

One advantage of the invention is that packets may be switched through the router at line rates without
5 requiring the storage of the packets in expensive high speed memory by providing a switching architecture that efficiently manages and routes packets through the switch.

Other advantages and features will be apparent
10 from the following description and claims.

Brief Description of the Drawings

Figure 1a and 1b are block diagrams of conventional router devices.

Figure 2a is a schematic block diagram of a data
15 routing system according to one embodiment of the present invention.

Figure 2b is a schematic block diagram of a router according to one embodiment of the present invention.

Figure 3 is a schematic block diagram of an input
20 port according to one embodiment of the present invention.

Figure 4a is a schematic block diagram of a router including the timing and ordering of transfers from the input port to input switch according to one embodiment of
25 the present invention.

Figure 4b is a data structure for a cell transferred between an input port and an input switch according to one embodiment of the present invention.

Figure 5a is a schematic block diagram of an input
30 switch according to one embodiment of the present invention.

- 6 -

Figure 5b is a schematic block diagram of a router including the timing and ordering of transfers from the input port to memory according to one embodiment of the present invention.

5 Figure 6 is a data structure for a route request stored in a key buffer according to one embodiment of the present invention.

Figure 7 is a data structure for an indirect cell according to one embodiment of the present invention.

10 Figure 8 is a data structure for a cell transferred between the input switch and a memory bank according to one embodiment of the present invention.

Figure 9 is a schematic block diagram of a reservation table according to one embodiment of the present invention.

Figure 10 is a flow diagram of a process of loading a reservation table according to one embodiment of the present invention.

20 Figure 11a is a schematic block diagram of a memory according to one embodiment of the present invention.

Figure 11b is a schematic block diagram of a router including the timing and ordering of transfers from the input port to memory according to one embodiment of the present invention.

Figure 12 is data structure for a cell output from a memory bank to output switch according to one embodiment of the present invention.

30 Figure 13 is a schematic block diagram of a controller according to one embodiment of the present invention.

Figure 14 is a data structure for an output request transferred from the controller to the output

- 7 -

switch according to one embodiment of the present invention.

Figure 15 is a schematic block diagram of an output switch according to one embodiment of the present invention.

Figure 16 is a data structure for a cell transferred from the output switch to an output port according to one embodiment of the present invention.

Figure 17 is a schematic block diagram for an output port according to one embodiment of the present invention.

Figure 18 is a flow diagram for a process of routing a packet through a router according to one embodiment of the present invention.

Figure 19 is a schematic block diagram of a router according to one embodiment of the present invention including sequence flow information for following the flow of operations performed by the router in routing a packet from an input port to its appropriate output port.

Figure 20 is a schematic block diagram of a multi-function port according to one embodiment of the present invention.

Figure 21 is a schematic block diagram of a router including multi-function port according to one embodiment of the present invention.

Detailed Description

Referring to Figure 2a, in a packet switching system, a source 10 is connected to one or more routers 20 for transmitting packets to one or more destinations 30. Each router includes a plurality of ports that are connected to various sources and destinations. A packet

- 8 -

from source 10 may pass through more than one router 20 prior to arriving at its destination.

Referring to Figure 2b, each router 20 includes an input switch 100, an output switch 102, a memory 104 including one or more memory banks 105, a controller 106 and a plurality of input and output ports 107 and 108, respectively. Associated with the controller 106 is controller memory 109 for storing a routing table. Input switch 100 is connected to each input port 107, while output switch 102 is connected to each output port 108 in router 20. In one embodiment, router 20 includes eight input and output ports 107 and 108, respectively.

In operation, packets are received at an input port 107, transferred to input switch 100 and stored temporarily in memory 104. When the packet is received by switch 100, a key is read from the first data block in the packet and transferred to controller 106. The key contains destination information which is derived from the header field associated with the first block of data in a packet and other information (source ID, flow ID, etc.).

A route look-up engine 110 in controller 106 performs a trie based search based on the key information and returns a result which includes the output port associated with the destination. The result is coupled with other information (source ID, flow ID, packet length, etc.) for routing the packet through router 20 and provided as a notification from controller 106 to output switch 102. Output switch 102 transfers the notification to the identified output port 108. Upon receiving the notification information, the output port 108 initiates the transfer of the packet from memory 104

- 9 -

through output switch 102 back to the appropriate output port 108.

Referring to Figure 3, each input port 107 includes a line input interface 300, a data handler 304 and a cell output port 306. Packets are received at line input interface 300. As the packets are received, data handler 302 divides the packets received into fixed lengths cells. In one embodiment of the present invention, the length of each cell is 80 bytes, with 16 bytes of internal header (control information) and 64 bytes of cell data. As the data handler divides the incoming packet into fixed length cells, it synchronously outputs the cells to input switch 100 through cell output port 306.

Referring now to Figure 4a, a single cell 450 is transferred from an input port 107 to input switch 100 at each cell slot "T". For a given cell slot "T", input port 107 receives a total of "N" cells, where "N" is equal to the number of input ports.

The data format for each cell 400 transferred from an input port 107 to input switch 100 includes an internal header 402 and a cell data field 404 as is shown in Figure 4b. The internal header 402 includes a type field 406, stream field 408, and packet header fields 410.

The type field 406 indicates the type of cell to be transferred from the input port. At each cell slot (20 clock cycles in one embodiment), an input port may transfer either a data cell, an indirect cell placeholder, or a delayed indirect cell placeholder. Data cells contain data associated with an incoming packet. An indirect cell placeholder is an empty cell, and is used in conjunction with indirect addressing for

- 10 -

the storage of the cells in the memory 104. Delayed indirect cell placeholders arise when a data stream that requires indirect addressing terminates at a time prior to the designated time for writing the last indirect addressing cell associated with the data stream to memory 104. The generation and operation of indirect placeholders and delayed indirect placeholders will be discussed in greater detail below in conjunction with Figure 7.

10 Stream field 408 indicates the stream to which the cell data belongs. In one embodiment of the present invention, each input port is capable of handling up to sixteen separate streams of data at a time.

Packet header field 410 contains header
15 information associated with a given packet and includes start offset information, packet length and interface index information.

Referring to Figure 5a, input switch 100 includes a round robin data handler 500, one or more input port
20 interfaces (501-0 through 501-7, one for each input port 107), one or more memory interfaces 502 (502-0 through 502-7, one associated with each memory bank), a like plurality of pointers 504 (504-0 through 504-7), an output processor 505, one or more output port interfaces
25 506 (506-0 through 506-7, one for each output port 108), a reservation table 508, an indirect cell processor 510, controller interface 512 and read controller 516.

Round robin data handler 500 receives cells from each input port and transfers them to output processor
30 505 for output to an appropriate memory bank 105 in memory 104. Round robin data handler 500 services the inputs (cells) received on input port interfaces 501 in a round robin, time division multiplexed manner. That is,

- 11 -

for a given cell slot, one cell from each input port is received at the round robin data handler 500 and subsequently transferred to output processor 505 for transfer at the next cell slot to a memory bank 105 in memory 104. At the next time cell slot, data handler 500 transfers the next cell received from the same input port to output processor 505 for transfer to a different memory bank. In one embodiment, the next cell received is transferred to the next memory bank (next in numerical order) in the memory array. Alternatively, another time dependent permutation may be used to control the transfer of successive cells from the same input port.

Referring to Figure 5b, the timing and ordering of transfers from the input to memory is shown. For the purposes of this example, a sequence of cells is depicted on each transmission line. For the purposes of this example only each transmission line is considered to be very long and contains data associated with two or more cells. In operation, the transmission lines are short and multiple cells are not present on a transmission line at a given time. At cell slot T4 a series of cells 450-0 through 450-7 are transferred down transmission lines 458, one from each input port 107 to input switch 100. At cell slot T3 (one just prior in time to cell slot T4) a series of cells 452-0 through 452-7 are transferred down transmission lines 458, one from each input port 107 to input switch 100.

Round robin data handler 500 and output processor 505 within the input switch 100 transfer cells out to memory 104 on transmission lines 460. As can be seen at cell slot T2, output processor 505 outputs one cell 454-B₀ to 454-B₇ to each memory bank in a single cell slot. The "B_x" designator indicates the input port from which the

- 12 -

particular cell was received. One cell from each input port is written to memory 104 per cell slot. At time period T1 (one cell slot prior to cell slot T2), again one cell (456-B₀ to 456-B₇) is written to each memory bank. Round robin data handler 500 time division multiplexes the transfers to output processor 505 such that consecutive cells from the same input port are written to consecutive memory banks 105 in memory 104.

Referring again to Figure 5a, pointer 504 indicates the location in an associated memory bank to which the next cell will be written. Output processor 505 writes a cell to a memory location in a particular memory bank based on the next available address in the bank as is indicated by the associated pointer 504.

Round robin data handler 500 includes a key reading engine 514 for determining the key information associated with a first cell in a packet and a linking engine 515 for linking cells in the same packet.

The process of reading key information is known in the art. After the key is determined for a given packet, it is stored temporarily in key buffer 516 in input switch 100 until the entire packet has been stored in memory 104. The data structure for entries 600 in the key buffer 516 is shown in Figure 6. Each entry 600 includes a key 602, full address 604, offsets 606 and an indirect cell indicator 608.

Linking engine 515 determines the starting address (full address) in memory for the first cell in a given packet. The starting address includes the bank number in memory 104 (the bank number which is assigned to store the cell by round robin data handler 500) and the first available address location in the designated bank (as is indicated by the associated pointer 504). The starting

- 13 -

address (full address 604) is stored in key buffer 516 along with the associated key 602 for the packet. When the next cell associated with the same packet arrives at switch 100, an offset 606 associated with the offset at which the cell is to be written (relative to the full address) is computed and stored in key buffer 516. In one embodiment of the present invention, up to four offsets 606 are stored. Each offset address is computed based on the relative offset in memory between the location of the last cell in memory and the value of the pointer 504 associated with the current memory bank which is to be written.

If more than five data cells are included in a packet, then the indirect cell indicator is set, and the last offset indicates the address in memory where the first indirect cell associated with the packet is stored. Indirect cells will be described in greater detail below in conjunction with Figure 7. After the packet has been stored in memory, the associated entry in key buffer 516 (a route look-up request) is forwarded via the controller interface 512 to the controller 106 for processing. Alternatively, the key may be transferred after the first five cells have been stored in memory.

The linking or threading of cells for the same packet is performed by using the offsets described above and indirect cells. Offsets are used to link the first 5 cells in a packet, while indirect cells are used to link the remaining cells in a packet. In one embodiment, if a cell contains 5 cells or less, no indirect cells are required to be used. Indirect cell processor 510 performs the linking of cells in memory for a given packet. Indirect cell processor 510 generates indirect cells for storage in memory 104. Indirect cells contain

- 14 -

offset information associated with the relative offset in memory space between contiguous cells in the packet.

Indirect cell processor includes indirect cell memory 520 for storing indirect cell data during the formation of
5 indirect cells.

Referring now to Figure 7, the data structure for an indirect cell 700 includes a linking field 702, a plurality of offset fields 704, and a last field 706. Linking field 702, when not set, indicates the current
10 indirect cell is the last cell in the chain of indirect cells for a given packet. If set, then more indirect cells exist for the given packet. If more indirect cells exist, then last field 606 indicates the offset to the location in memory of the next indirect cell associated
15 with the packet. In one embodiment of the present invention, each indirect cell contains 56 offset data blocks for linking 56 cells in memory.

As was described above, when a packet is received, the linking engine processes the first five cells and
20 stores linking information in the form of a start address and four offsets in key buffer 516. In the event more than five cells are contained within a packet, the indirect cell processor takes over for the linking engine and computes the offsets associated with the locations in
25 memory where the remaining cells in the packet are stored. Round robin processor 500 passes cells to the output processor 505 for transfer to an associated memory bank in memory 104. Round robin processor 500 enables the indirect cell processor when the packet being
30 processed contains more than 5 cells (based on header information included within the first cell). At the time for writing the fifth cell to memory, indirect cell processor 510 stores in indirect cell memory 520 the

- 15 -

address (the "indirect cell address") associated with the location in memory at which the fifth cell would have been written if it had been the last cell in the packet. The indirect cell address indicates the location in
5 memory where the indirect cell is to be written when it is full (or when the last cell of the packet is processed).

When a indirect cell is full (having stored offsets in all available locations except the last field
10 706), then the indirect cell processor stores the offset associated with the location in memory where the next indirect cell is located in the last field 606. Thereafter, the full indirect cell is written to its appropriate place in memory. The writing of the indirect
15 cell to memory coincides with the receipt of an indirect cell placeholder by the input switch 100 from the associated input port 107. This process continues until the last cell in a packet is stored in memory. At that time, the last indirect cell is written to memory, and
20 the associated entry 600 from the key buffer 516 is transferred to the controller 106 for processing.

As often will be the case, the last cell of a packet will not coincide with the timing required to write the completed indirect cell immediately into
25 memory. This is because packet length is completely arbitrary. The end of a packet will likely not coincide with the filing of an entire indirect cell. When a packet has completed (all cells have been received by the input switch) and the last entry in the indirect cell is
30 written, the indirect cell is free to be written to memory. However, the writing will be delayed until the proper time, hence the term delayed indirect cell. A delayed indirect cell is a indirect cell that is the last

- 16 -

indirect cell associated with a packet. It is delayed, because it is written to memory after the rest of the packet has been written to memory. The timing of the write to memory is dictated by the address which is reserved for the indirect cell. As was described above, at the time for the creation of an indirect cell, its position in memory is reserved. The delayed indirect cell will be written to memory at the next time slot available for the particular input port to write to the particular memory bank after the packet has been completed. The timing of the write to memory of delayed indirect cells coincides with the receipt of a delayed indirect placeholder from the appropriate input port 107.

Read controller 517 controls the transfer of read request signals flowing from input switch 100 out memory interface 502 to the individual memory banks in memory 104. Read controller 517 receives read requests from each output port through output port interfaces 506. The format of each request includes source identification (output port) and a full address in memory which is to be read. At each cell slot, each output port may generate a read request for processing by switch 100 to read a memory location in memory 104, resulting in the reading of a cell (a read reply) from a memory bank 105 (on a subsequent cell slot) to output switch 102.

The data structure of a cell transferred from input switch 100 (via the output processor 505) to a memory bank 105 in memory 104 is shown in Figure 8. At each cell slot, output processor 505 generates a cell 800 which includes a read request source field 802, read address 804, write address 806 and data field (cell data received from input port 107) 808. The read request source field 802 indicates the output port requesting

- 17 -

the read (destination output port 108). Output processor 505 receives read requests from read controller 517 and bundles the read request with any write request received from round robin data handler 500 destined for the same
5 memory bank. At each cell slot, output processor 505 provides a cell 800 which may include a write and read request to each memory bank 105 in memory 104.

Read controller 517 loads a reservation table 508 as requests to transfer packets are received from the
10 various output ports 108. The reservation table is loaded such that at every cell slot a single read request is generated for each bank of memory 105. Referring now to Figure 9, reservation table 508 includes a plurality of columns 900, one for each memory bank 105 in memory
15 104, a plurality of rows 902, placeholders 904 and loaded entries 906. Each row represents a set of read requests (one per memory bank) to be generated on a single cell slot. Each row includes a single entry for each output port 108. At each cell slot, each output port is capable
20 of requesting a read from a single memory bank 105 in memory 104. Associated with reservation table 508 is a read pointer 908. The pointer points to the next row in the reservation table to be read. Rows ahead of the read pointer correspond to requests that will be queued at a
25 later cell slot time. In one embodiment, the pointer moves at least one row in each cell slot time.

Loaded entries 906 reflect read requests to be performed as a result of reservation requests received from output switch 102. Placeholders 904 represent
30 available read requests which have not as of yet been requested. At each cell slot, the read controller 517 performs three functions: loading entries in the reservation table at the first available location in the

- 18 -

table (after the read pointer) outputting the last row as read requests to the output processor 505; and refreshing the table, moving out the last row, incrementing the rows and creating a new row at the top of the table. The
5 number of rows in the reservation table must be as large as the product of the latency in processing read requests times the number of banks. In one embodiment, 48 rows are included in reservation table 508 reflecting a system including six cell slots of latency and eight memory
10 banks.

At initialization, reservation table 508 contains placeholders 904 in all of the rows 902. Placeholders 904 are locations in the reservation table which have not been loaded. As read requests are processed by the read
15 processor, certain ones of the placeholders 904 are converted to loaded entries 906 based on the read requests. Loaded entries 906 include a read request address.

Referring now to Figure 10, the process of loading
20 the reservation table includes receiving a read request (full address) from an output port (1000). The read controller decodes the read request to determine the column (based on the memory bank to be read from) in the reservation table to search (1002). The read processor
25 searches, starting at the bottom of the reservation table, for the first placeholder associated with the output port that generated the read request (1004). The read processor transforms the placeholder 904 to a loaded entry 906 by writing the full address of the read request
30 at the location (1006). The process repeats for each read request received by the read controller (1008).

Referring now to Figure 11, memory 104 includes a plurality of memory banks 105. Each memory bank includes

- 19 -

a input port 1102 and output port 1104. At each cell slot, each memory bank receives at most one write and one read request via input port 1102. The write requests are associated with cells received from input ports 107.

5 Read requests reflect a request for cell data to be transferred from a memory bank 105 to output switch 102. The data structure associated with the cells written from memory 104 to output switch 102 is shown in Figure 12. Each cell 1200 includes an output port identifier 1202
10 and cell data 1204.

In one embodiment, the memory is divided into a plurality of banks where the number of memory banks is equal to the number of input and output ports. A one to one relationship exists between input ports, output ports
15 and memory banks. In this embodiment, the transfer of cells from the input switch 100 to memory 104 is performed in a time division multiplex fashion. That is, consecutive cells from a given input port are directed to different memory destination locations. At each time
20 period (cell slot), the input switch transfers to memory a single cell from each input port (as available) into memory. At a next time $T+1$ the input switch transfers again a single cell from each input port into memory. Successive entries from the same input port are written
25 to different memory banks 105 in memory 104.

Referring now to Figure 13, controller 106 includes controller memory 109, route look-up engine 110, input switch interface 1300, and output switch interface 1302. Controller 106 receives a route look-up request
30 from input switch 100 at the input switch interface 1300. In one embodiment of the present invention a plurality of route look-up engines 110 are included in controller 106, each receiving look-up requests in round-robin fashion so

- 20 -

as to speed the routing process. In one embodiment, controller memory 109 is a four bank static random access memory (SRAM) that requires thirty two route look-up engines 110 to service at full bandwidth. The matching
5 of keys retrieved from a packet in order to determine a best match route through the router is described in greater detail in co-pending patent application entitled "HIGH SPEED VARIABLE LENGTH BEST MATCH LOOK-UP IN A SWITCHING DEVICE", filed on December 16, 1996, by
10 Ferguson et al., serial number 08/767,576, which is hereby expressly incorporated by reference.

The route look-up engine servicing the route look-up request performs a best match look-up and outputs notification through output switch interface 1302 to
15 output switch 102. The notification includes a result which indicates the output port to be used in the transfer of the packet to its destination.

Referring now to Figure 14, the data structure associated with the notification outputted by the
20 controller 106 to the output switch 102 is shown. The data structure 1400 for the notification includes a mask 1402, a next hop index pointer 1404, full address 1406, offsets 1408 and packet length 1410.

The mask field 1402 is used to indicate which
25 output port connected to output switch 102 is to transfer the packet. In one embodiment, the notification may be sent to more than one output port resulting in the broadcast of the associated packet.

Associated with each output port 108 is a memory.
30 The next hop index pointer points to a location in the memory. The memory is used to store media header information associated with a particular type of packet transfer. Next hop addresses and media headers will be

- 21 -

described in greater detail below in association with output port 108.

The full address 1406 indicates the starting address in memory where the first cell in the packet is
5 stored. As was described above, offsets 1408 provide linking information for retrieving cells or an indirect cell associated with the packet.

Referring now to Figure 15, output switch includes a controller interface 1500, one or more memory inputs
10 1502 (1502-0 through 1502-7, one for each memory bank), one or more outputs 1504 (1504-0 through 1504-7, one for each output port), a result processor 1506 and an output processor 1508. Output switch 102 performs four
15 results, receive output results, process output results, receive cells from memory and output cells to output ports.

Cells from memory are received at memory inputs 1502 and transferred to output processor 1508. Output processor 1508 decodes the destination output port from
20 the cell information received from memory and transfers the cell data to the appropriate outputs 1502. At each cell slot, output switch 102 receives a cell for processing from each bank in memory 104.

Output switch 102 receives notification from
25 controller 106 on controller interface 1500. Result processor 1506 decodes the result (route) and determines which output port(s) 108 is (are) to receive the route data. Based on mask 1402 in the notification, result processor 1506 transfers the notification to output
30 processor 1508 for transfer to each respective output port 108 so indicated. At each cell slot, output processor 1508 provides (via outputs 1504) a route to each output port 108.

- 22 -

The data structure associated with the data transferred from output processor 1508 to output ports 108 is shown in Figure 16. A cell 1600 includes a header 1602 and data field 1604. The header 1602 includes
5 memory bank source information 1606 and route information 1608. The memory bank source information includes a source identifier for indicating which memory bank provided the cell in data field 1604. Route information 1608 contains data from the notification including a next
10 hop index, packet length, full address and offsets.

Referring now to Figure 17, each output port 108 includes an output switch interface 1700, an input switch interface 1702, a buffer 1704, an output request processor 1706, an line output interface 1708, storage
15 device (memory) 1710, output buffer 1712 and output formatter 1714.

Output ports 108 receive notification that a packet is to be processed by cells 1600 received at the output switch interface 1700. The output request
20 processor 1706 stores the request in buffer 1704 and subsequently generates a read request to input switch 100 associated with the first address in memory where the packet is stored. The output request processor 1706 generates the first read request based on the full
25 address received from output switch 102. Thereafter subsequent read requests are generated for transmission to the input switch based on the offset information provided in the request (from cell 1600) or indirect cells (as will be described below).

30 If the packet length, as determined from the route information provided with the cell 1600, is greater than five (5) cells, then the output request processor first requests the transfer (read from memory) of the first

- 23 -

indirect cell associated with the packet. This is accomplished by computing the address of the indirect cell based on the full address and the offsets provided in cell 1600. After the indirect cell request is
5 generated, the output request processor generates read requests for the remaining cells in the packet based on the full address and the offsets provided in cell 1600. Upon receipt of a indirect cell from the output switch 102, output request processor continues to generate read
10 requests for the remaining cells in the packet based on the offset information contained within the indirect cell.

Subsequent indirect cells are retrieved in a similar fashion. That is, at the time for reading the
15 next indirect cell, the address of the next indirect cell is computed based on the last offset stored in the previous indirect cell. The timing of retrieving the indirect cells is accomplished such that no delays in the output stream are required. Each subsequent indirect
20 cell is retrieved prior to the end of the processing of the prior indirect cell. In this way, once the output stream is initialized, no buffering of data is required and no interruptions due to the latency associated with the retrieval process are experienced.

25 Output requests to an individual memory bank are processed strictly in order. That is, output port may track each request issued to a memory bank and is assured that the data received in response to a series of requests to the same memory bank will be strictly
30 delivered according to the sequence or pattern which they were issued. Output request processor keeps track of requests generated for each memory bank. Output buffer 712 includes a plurality of queues, two for each memory

- 24 -

bank (one request queue and one reply queue), that are used to order cells received from memory according to the particular stream to which they are assigned. The request queue contains a stream number and a read
5 address. When a request is issued to memory, the entry is removed from the request queue and the stream number portion is placed in the reply queue. When a reply is received, the entry at the head of the reply queue is removed and the reply is sent to the stream number
10 indicated by the stream number retrieved from the reply queue.

As cells are received back at the output port 108 (responsive to the read requests), they are stored in output buffer 1712. For given packet, the output port
15 stores the number of cells required to provide a streamed output. In one embodiment of the present invention, twelve cells are stored prior to beginning to output (stream data) from the output port. The selection of the number of cells for storage in output buffer 1712 is
20 based on the latency in the read process (number of clock cycles between a read request from an output port and the arrival of the cell associated with the read request to the output port).

Output formatter 1714 receives the cells from
25 output buffer 1712 and couples the data with media header information stored in memory 1710. Each request (notification) received from output switch 102 includes a next hop index. The next hop index indicates the starting address in memory 1710 of the media header
30 information associated with a given type of transmission (derived from the destination of the packet). Output formatter 1714 couples the cell data returned from memory with the appropriate media header to generate a proper

- 25 -

packet for transfer out of router 20 on the line output interface 1708.

Referring now to Figure 18, in a method of routing packets through a switch a packet is received at an input port (1800). The input port divides the packet into fixed length cells and transfers the cells to an input switch (1802). Input switch removes the key information from the first cell in a packet and stores it temporarily in a key buffer (1804). Thereafter the input switch routes the cells to memory banks in a time division multiplexed manner (1806). The input switch stores the first address associated with where the first cell is written in memory and computes offsets for each additional cell associated with the offset in memory for the next contiguous memory bank into which the next cell is written (1808). The input switch creates indirect cells to store linking information for the packet if the packet length exceeds five cells (1810). If the number of cells exceeds the number of available offsets in an indirect cell, then the old indirect cell is stored in memory and a new indirect cell is created and loaded based on the offsets calculated for each new cell received at the input switch.

When the packet (and its indirect cells if any) have been stored in memory, then the key, full address of the first cell and offset information is transferred as a look-up request to a controller (1814). The controller performs a best match look-up and generates a result of the look-up. The result which includes the destination port (output port), address, offset information and next hop index (1816). A notification including the result is transferred to the output switch for transfer to the appropriate output port (1818).

- 26 -

Upon receipt of a notification, the output port generates read requests a cell at a time to the input switch for the data associated with the packet (1820). The input switch issues the read requests in a time division multiplexed fashion generating a single request to each memory bank per cell slot (1822). When the memory bank receives the request from the input switch, the output port associated with the request and the cell data is written to the output switch (1824). Again, at each cell slot, the output switch transfers a single cell to each of the output ports. Upon receipt, the output port couples the cell data with media header information and streams the data to the destination (1826).

15 Alternative Embodiments

In one embodiment of the present invention, an output port, an input port and a memory bank are contained in a single device. The architecture of this multifunction port including memory is shown in Figure 19. Specifically, the multifunction port 1900 includes a memory bank 105, a line input interface 300, a data handler 302, a buffer 1704, output request processor 1706, a line output interface 1708, a storage device 1710, a FIFO 1712, output formatter 1714, an input switch interface 1902 and an output switch interface 1904.

The multifunction port is used in conjunction with the input switch, output switch and controller as is shown in Figure 20. The various piece components of the input port, output port and memory bank perform the identical functions described above. However, the combination of the devices into a single unit simplifies the interfaces between the components. Specifically, the

- 27 -

revised format for transfers between the multifunction port and the input switch is shown in Figure 21.

A cell 2100 transferred from a multifunction port 1900 to the input switch contains a cell header 2102 and
5 cell data 2104. Cell header 2102 includes a type field 406, stream field 408, and packet header fields 410 similar in function to those fields described above in reference to Figure 4. In addition, cell header 2102 includes a read request in the form of a output port
10 identifier 2106 and address 2108. Output port identifier 2106 identifies the output port which is sourcing the read request. Address 2108 indicates the address in memory 104 to be read.

The present invention has been described in terms
15 of specific embodiments, which are illustrative of the invention and not to be construed as limiting. Other embodiments are within the scope of the following claims.

What is claimed is:

- 28 -

1. A router for switching a data packet between a source and destination in a network comprising:

an input port including a data handler, the input port receiving the data packet from the source, the data
5 handler dividing the data packet into one or more fixed length cells;

an output port for routing the data packet to the destination;

a memory divided into a plurality of memory banks;
10 an input switch for receiving the fixed length cells from the input port and routing a single cell in a cell slot time span to each memory bank; and

an output switch for routing cells received from the memory to the output port.

15 2. The router of claim 1 where the input switch includes a linking engine for linking cells in the data packet to allow retrieval of the data packet from non-contiguous locations in the memory.

3. The router of claim 2 further including an
20 indirect cell generator for generating one or more indirect cells, the linking engine tracking the location in the memory where consecutive cells of the data packet are stored and providing an address in memory of each cell in the data packet for storage in indirect cells.

25 4. The router of claim 1 wherein the input switch time division multiplexes the writing of data packets to the memory such that consecutive cells from the input port are written to consecutive banks in the memory.

- 29 -

5. The router of claim 1 wherein the input switch includes a key reading engine for extracting key information from a first cell received at the input switch associated with the data packet, the router
5 further including a controller coupled to the input switch and receiving the key information therefrom, the controller for decoding destination information from the key information received from the input switch and outputting a notification defining a routing of the data
10 packet from the memory to the output port.

6. The router of claim 5 wherein the output port includes a result processor for receiving the notification from the controller and initiating a transfer of the data packet from the memory to the output
15 port.

7. The router of claim 1 wherein the input switch includes a reservation table for scheduling transfers from the memory to the output switch.

8. The router of claim 7 further including a
20 controller coupled to the input switch and the output switch for decoding destination information received from the input switch and outputting a notification to the output switch defining a routing of the data packet from the memory to the output port.

25

9. The router of claim 8 wherein the output switch routes the notification to the output port and thereafter the output port issues a request to the input switch to transfer the data packet from memory to the
30 output port through the output switch.

- 30 -

10. The router of claim 9 wherein the request from the output port is stored in the reservation table.

11. The router of claim 10 wherein requests to transfer cells from memory to the output switch are time domain multiplexed so that during one cell slot time span at most a single read request is issued to each bank in the memory for servicing.

12. The router of claim 9 wherein the memory outputs at most a single cell per bank in one cell slot time span.

13. A router for switching a data packet between a source and destination in a network comprising;
an input port for receiving a data packet from the source, the input port including a data handler for dividing the data packet into fixed length cells;
a memory divided into a plurality of memory banks;
an input switch including a linking engine, the input switch receiving at most a single cell from the input port in a cell slot time span and routing at most a single cell from the input port to the memory in a cell slot, the input switch time division multiplexes writing of cells to the memory such that consecutive cells from the input port are written to consecutive banks in the memory, the linking engine for linking cells in the data packet to allow retrieval of the data packet from non-contiguous locations in the memory;
a controller for decoding destination information associated with the data packet, the controller outputting a notification defining a routing of the data packet through the router;

- 31 -

an output port including a result processor for receiving the notification from the controller and initiating a transfer of the data packet from memory to the output port; and

5 an output switch for routing cells received from the memory to the output port.

14. A router for switching a data packet between a source and destination in a network comprising:

a plurality of input ports each including a data
10 handler, a first input port receiving the data packet from the source, the data handler of the first input port dividing the data packet into one or more fixed length cells;

a plurality of output ports at least one of which
15 is for routing the data packet to the destination;

a memory divided into a plurality of memory banks;

an input switch for receiving fixed length cells from one or more input ports and routing a single cell in a cell slot time span to each memory bank; and

20 an output switch for routing cells received from the memory to an appropriate output port.

15. The router of claim 14 where the input switch includes a linking engine for linking cells in the data packet to allow retrieval of the data packet from non-
25 contiguous locations in the memory.

16. The router of claim 15 further including a indirect cell generator for generating one or more indirect cells, the linking engine tracking the location in the memory where consecutive cells of the data packet

- 32 -

are stored and providing an address in memory of each cell in the data packet for storage in indirect cells.

17. The router of claim 14 wherein the input switch time division multiplexes the writing of cells to
5 the memory such that consecutive cells from any input port are written to different banks in the memory.

18. The router of claim 14 wherein the input switch includes a key reading engine for extracting key information from a first cell received at the input
10 switch associated with the data packet, the router further including a controller coupled to the input switch and receiving the key information therefrom, the controller for decoding destination information from the key information received from the input switch and
15 outputting a notification defining a routing of the data packet from the memory to an appropriate output port.

19. The router of claim 18 wherein the output port includes a result processor for receiving the notification from the controller and initiating a
20 transfer of the data packet from the memory to the appropriate output port.

20. The router of claim 14 wherein the input switch includes a reservation table for scheduling transfers from the memory to the output switch.

- 33 -

21. The router of claim 20 further including a controller coupled to the input switch and the output switch for decoding destination information received from the input switch and outputting a notification to the
5 output switch defining a routing of the data packet from the memory to the appropriate output port.

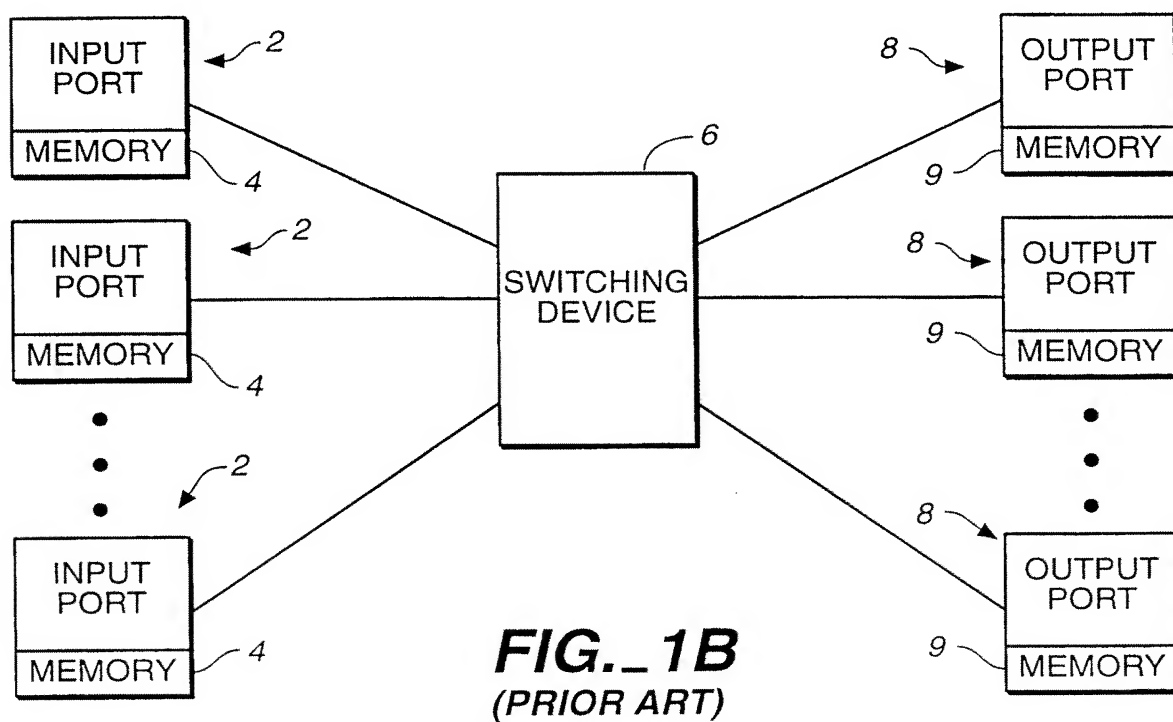
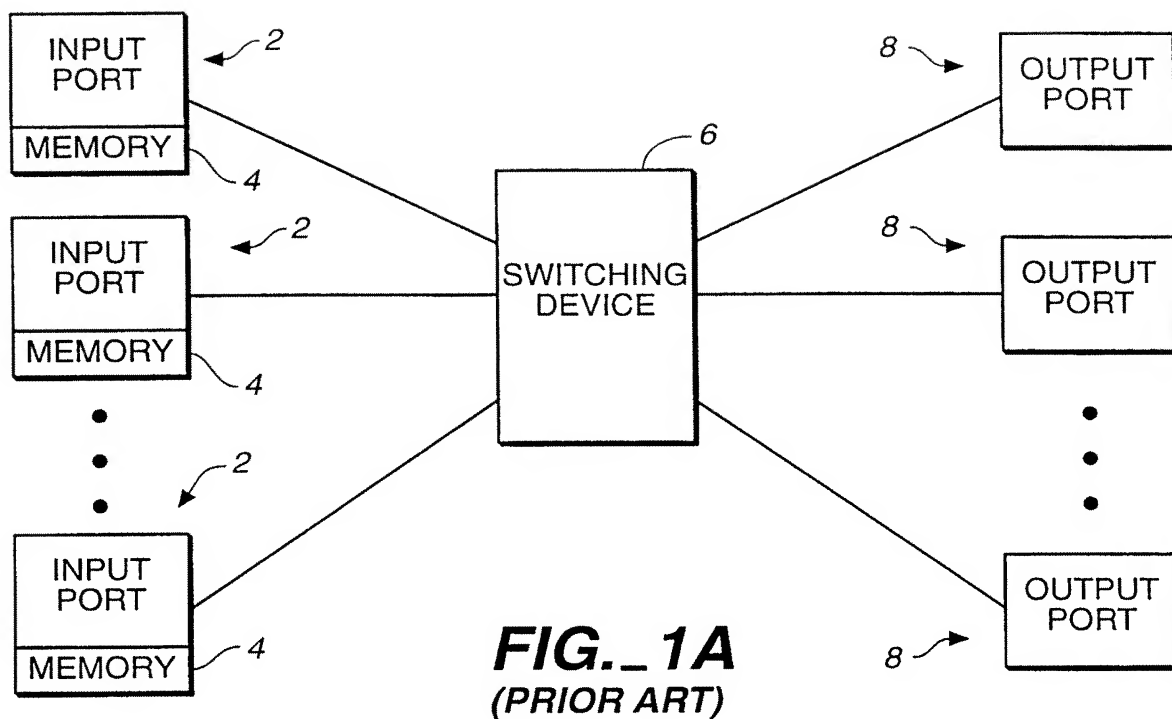
22. The router of claim 21 wherein the output switch routes the notification to the output port and
10 thereafter the output port issues a request to the input switch to transfer the data packet from memory to the appropriate output port through the output switch.

23. The router of claim 22 wherein the request
from
15 the output port is stored in the reservation table.

24. The router of claim 23 wherein requests to transfer cells from memory to the output switch are time domain multiplexed so that during one cell slot time span at most a single read request is issued to each bank in
20 the memory for servicing.

25. The router of claim 22 wherein the memory outputs at most a single cell per bank in one cell slot time span.

1 / 18



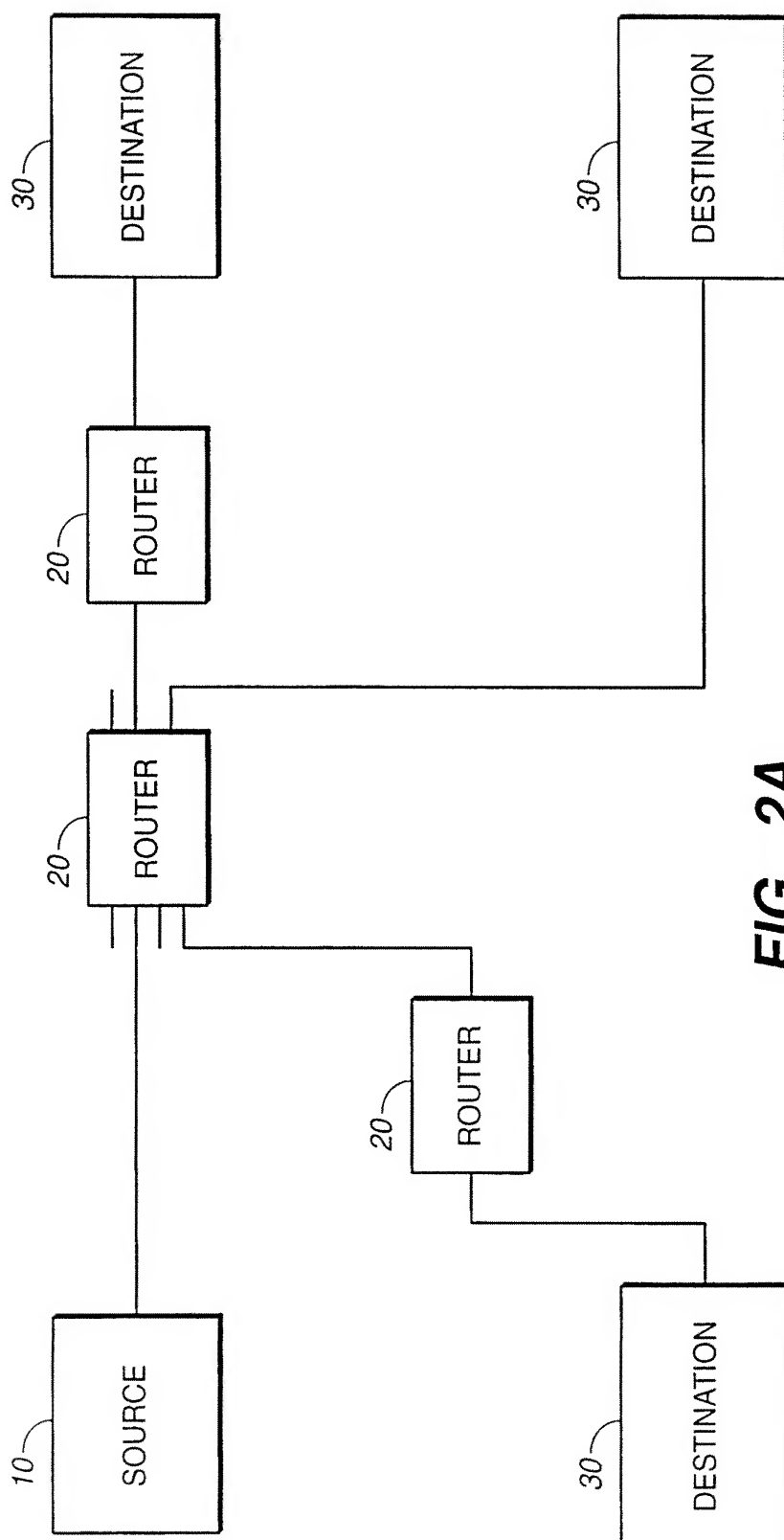


FIG. 2A

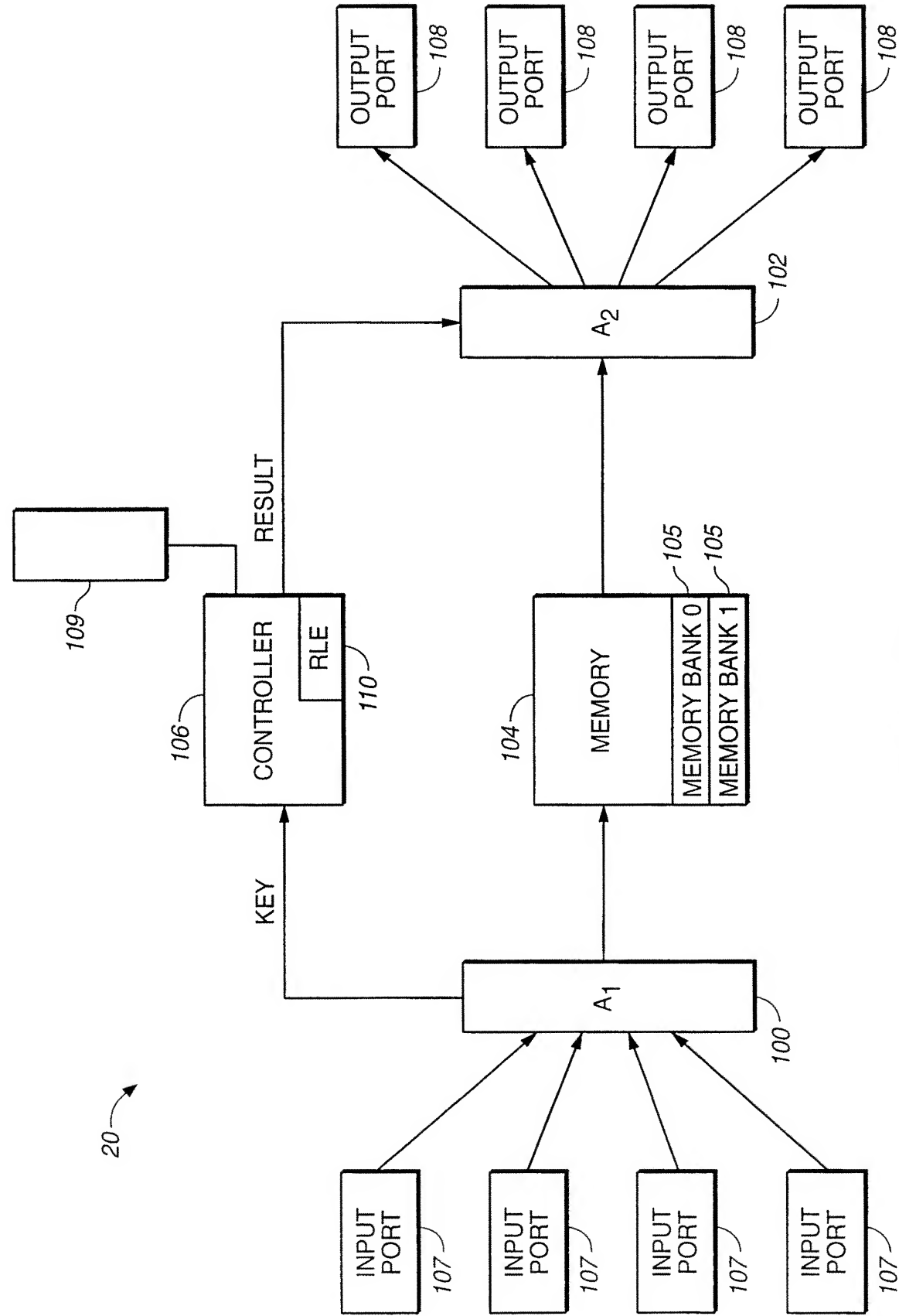


FIG. 2B

4 / 18

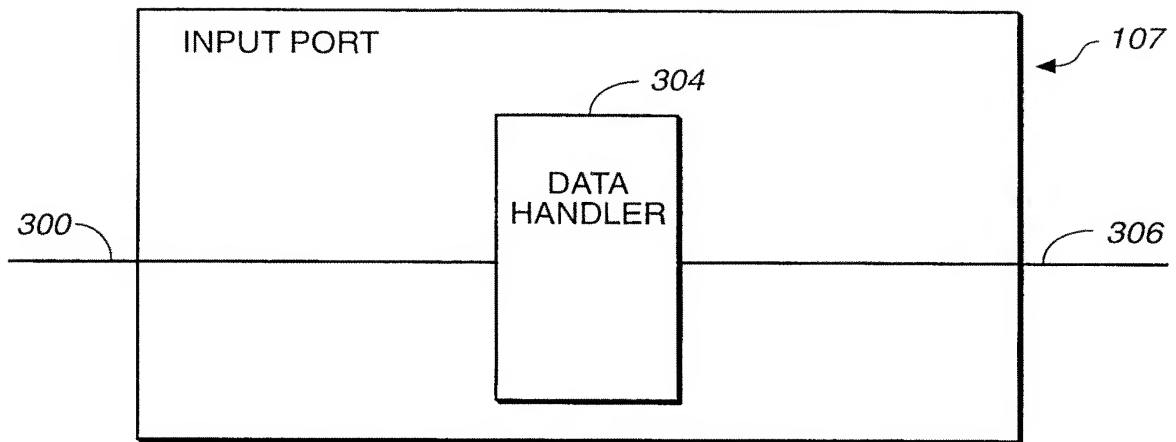


FIG._3

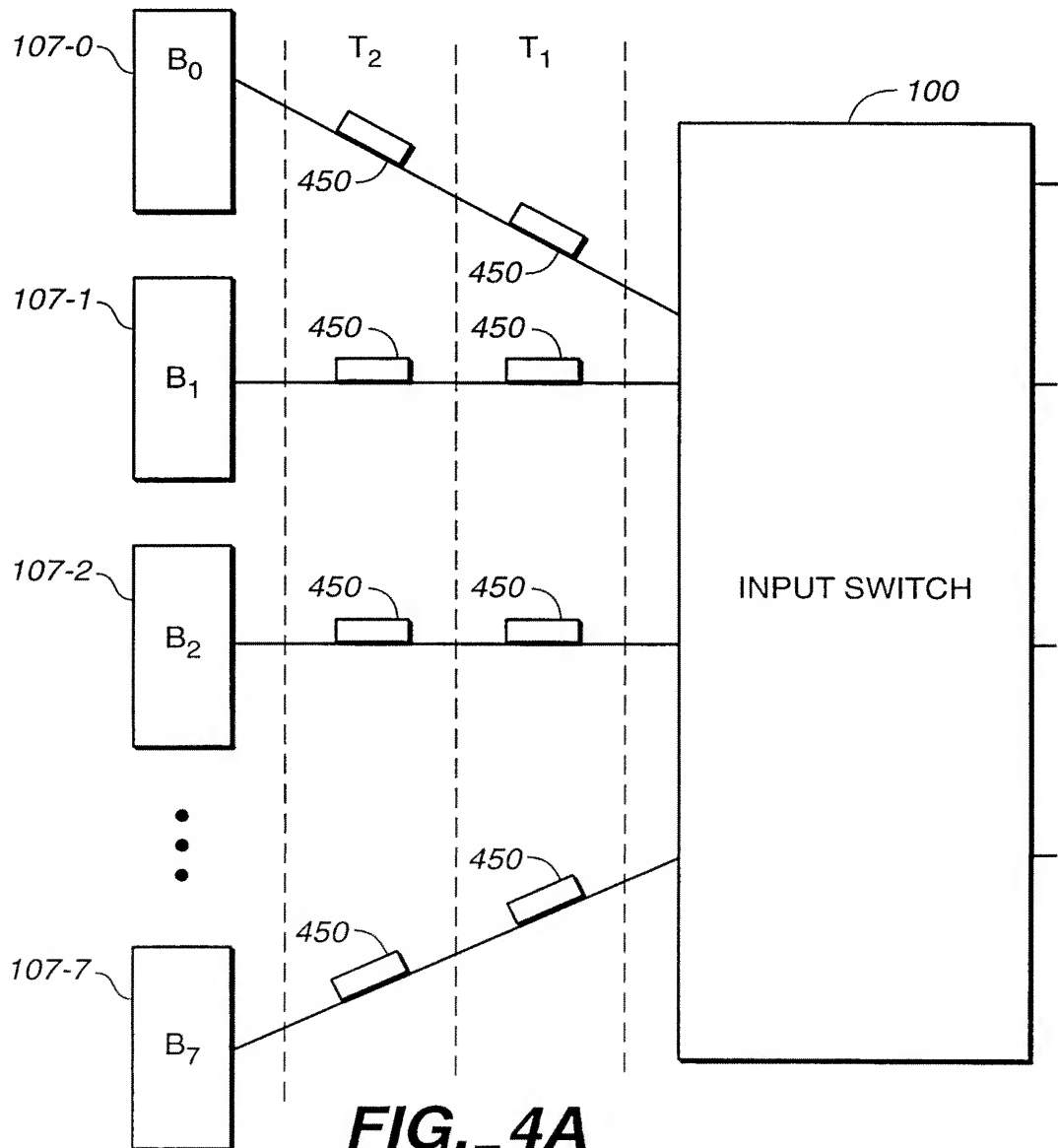


FIG._4A

5 / 18

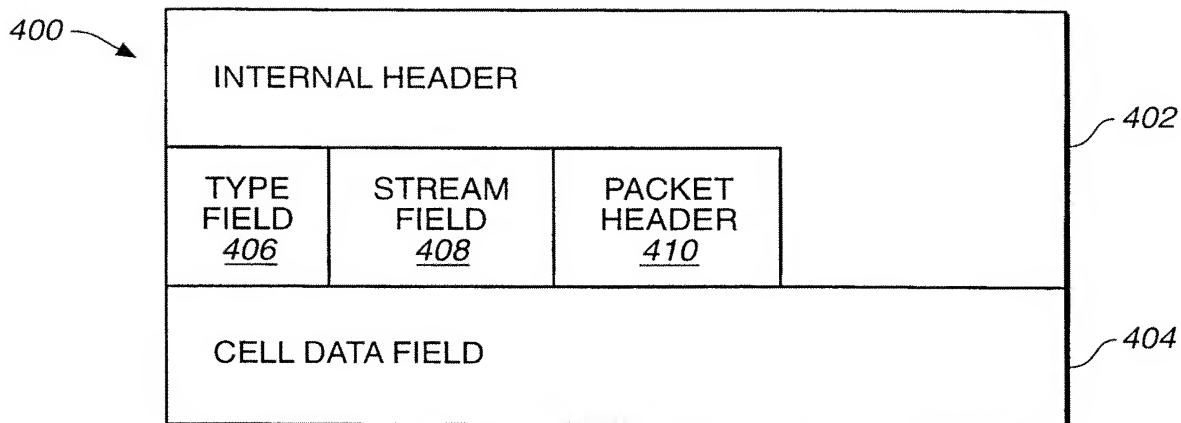


FIG. 4B

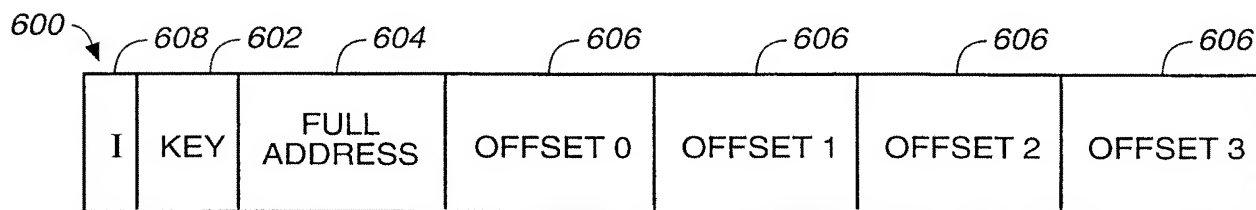


FIG. 6

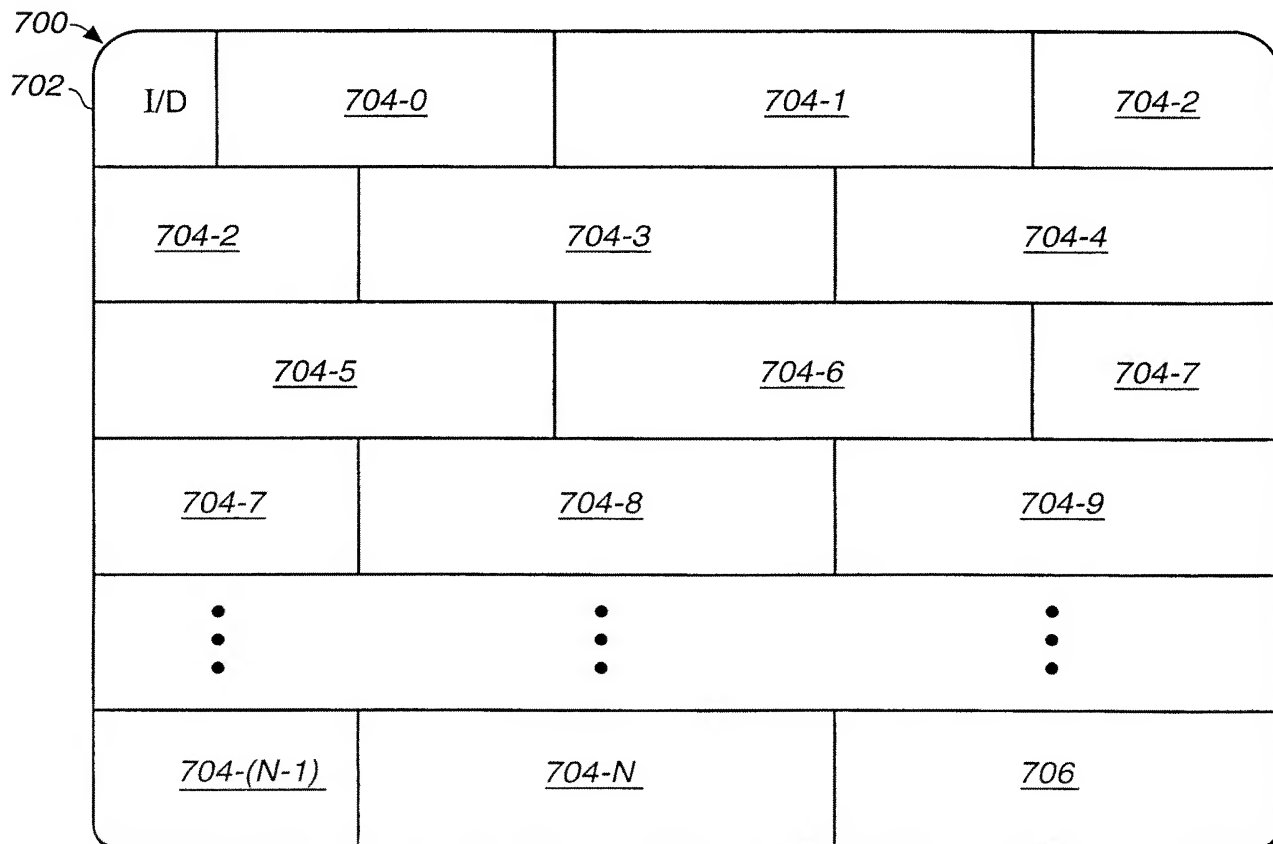


FIG. 7

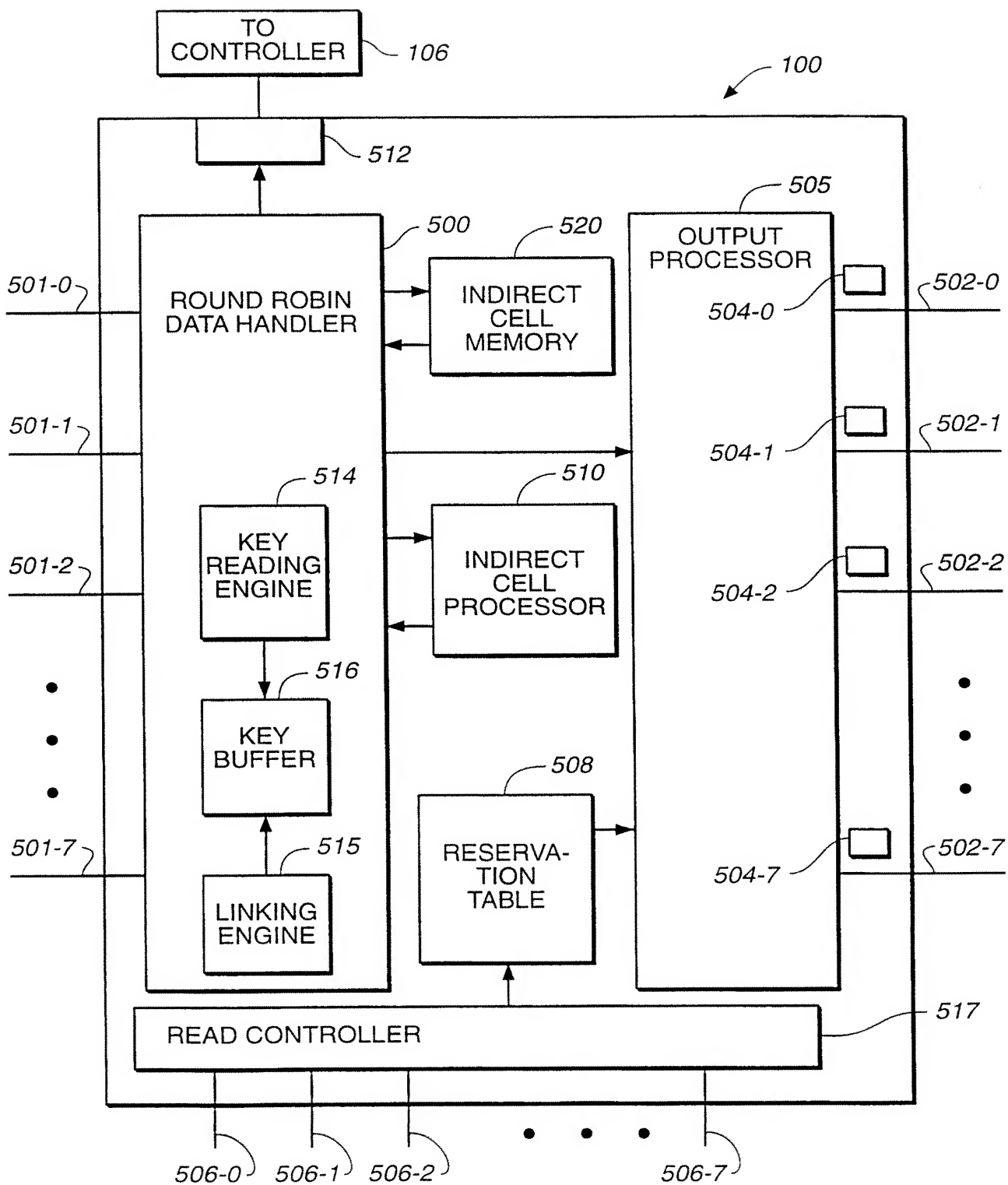
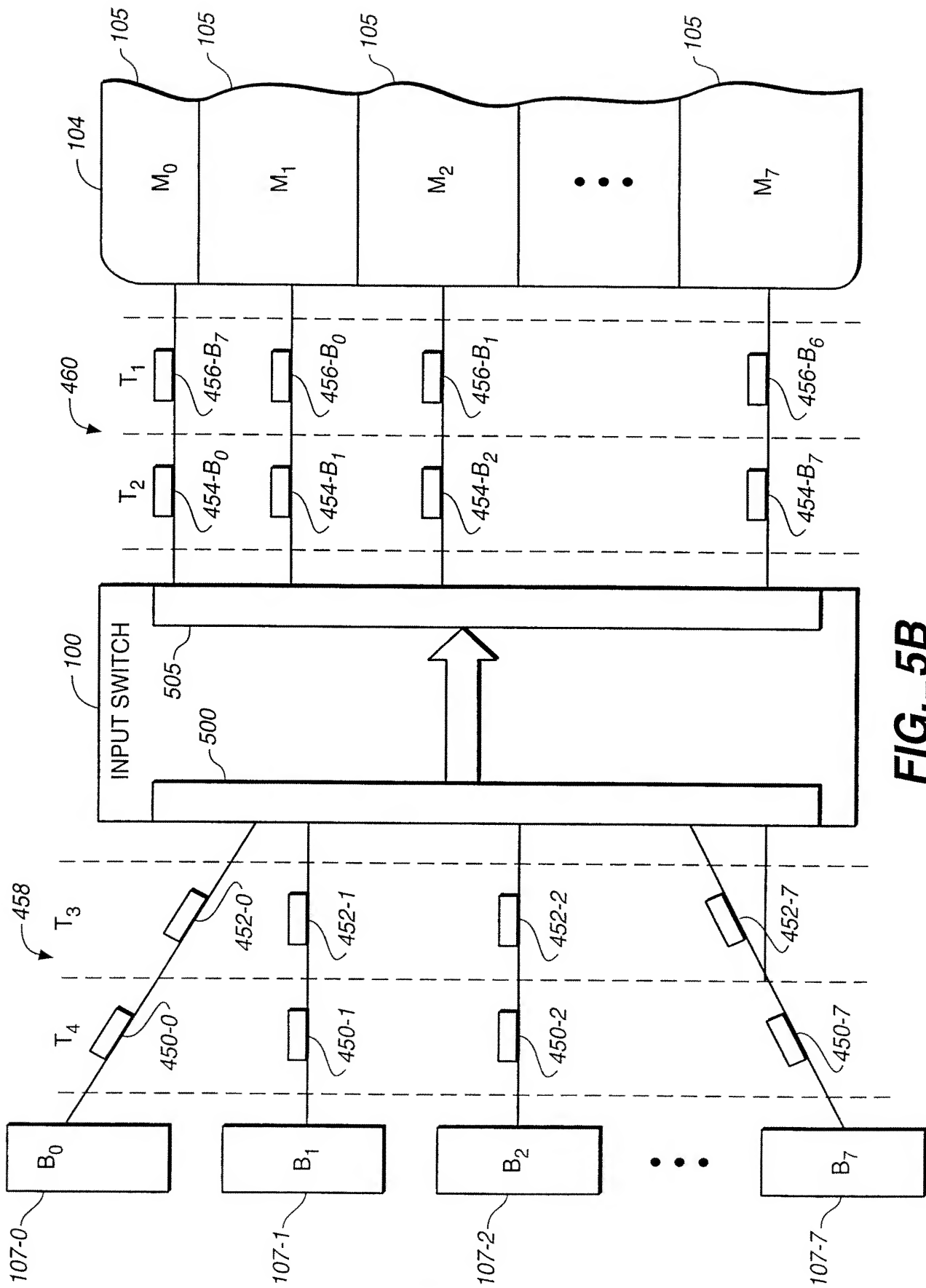
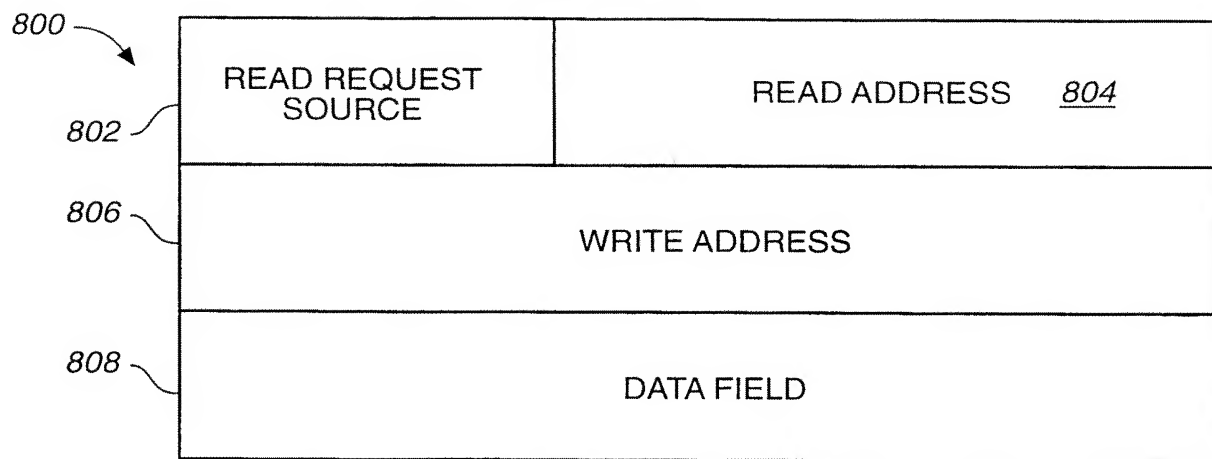
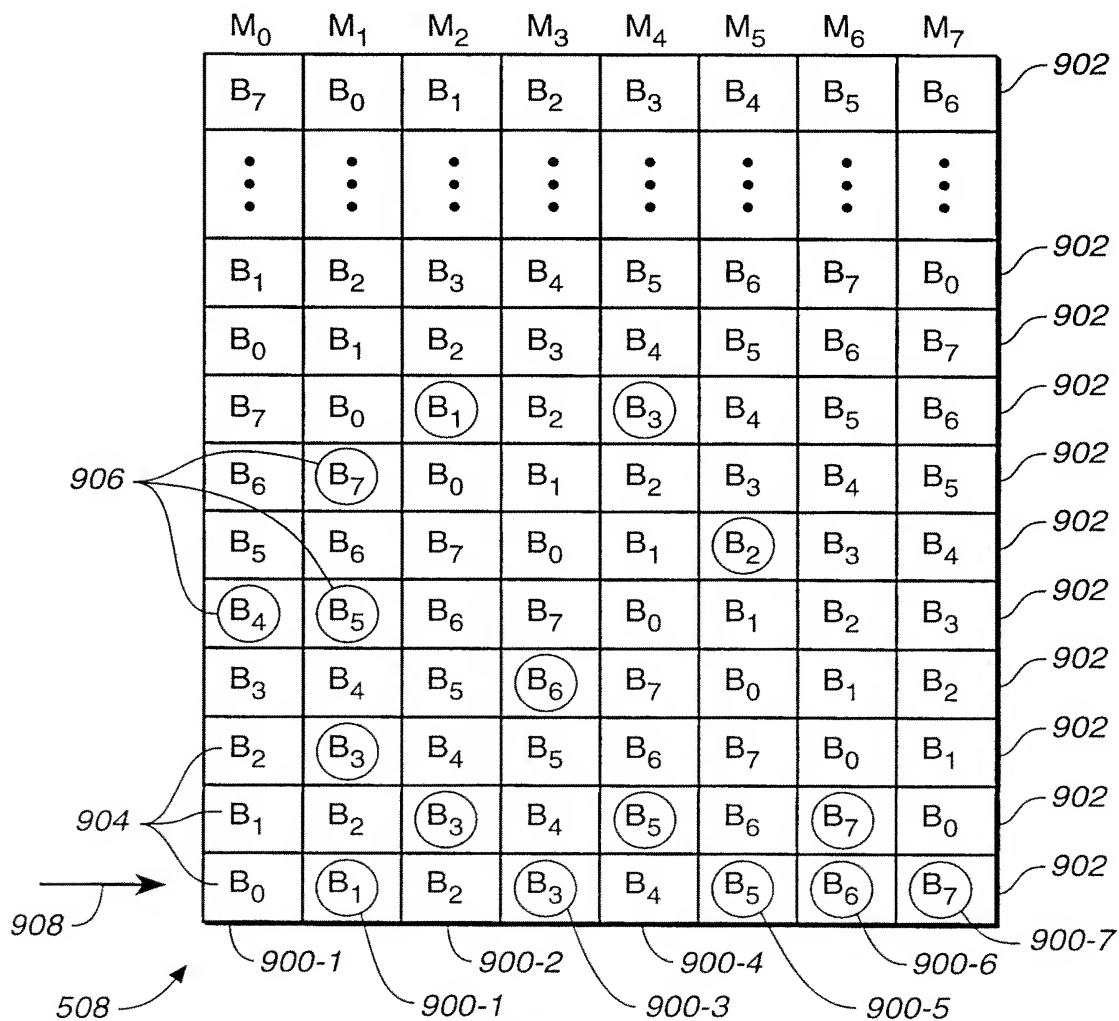


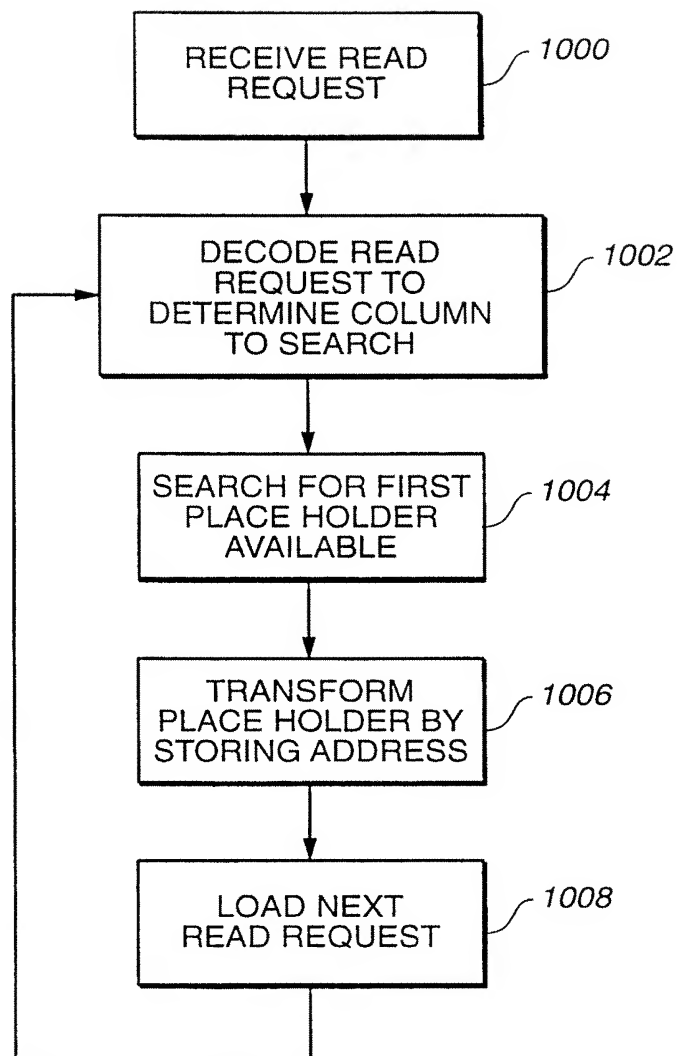
FIG. 5A



8 / 18

**FIG. 8****FIG. 9**

9 / 18

**FIG. 10**

10 / 18

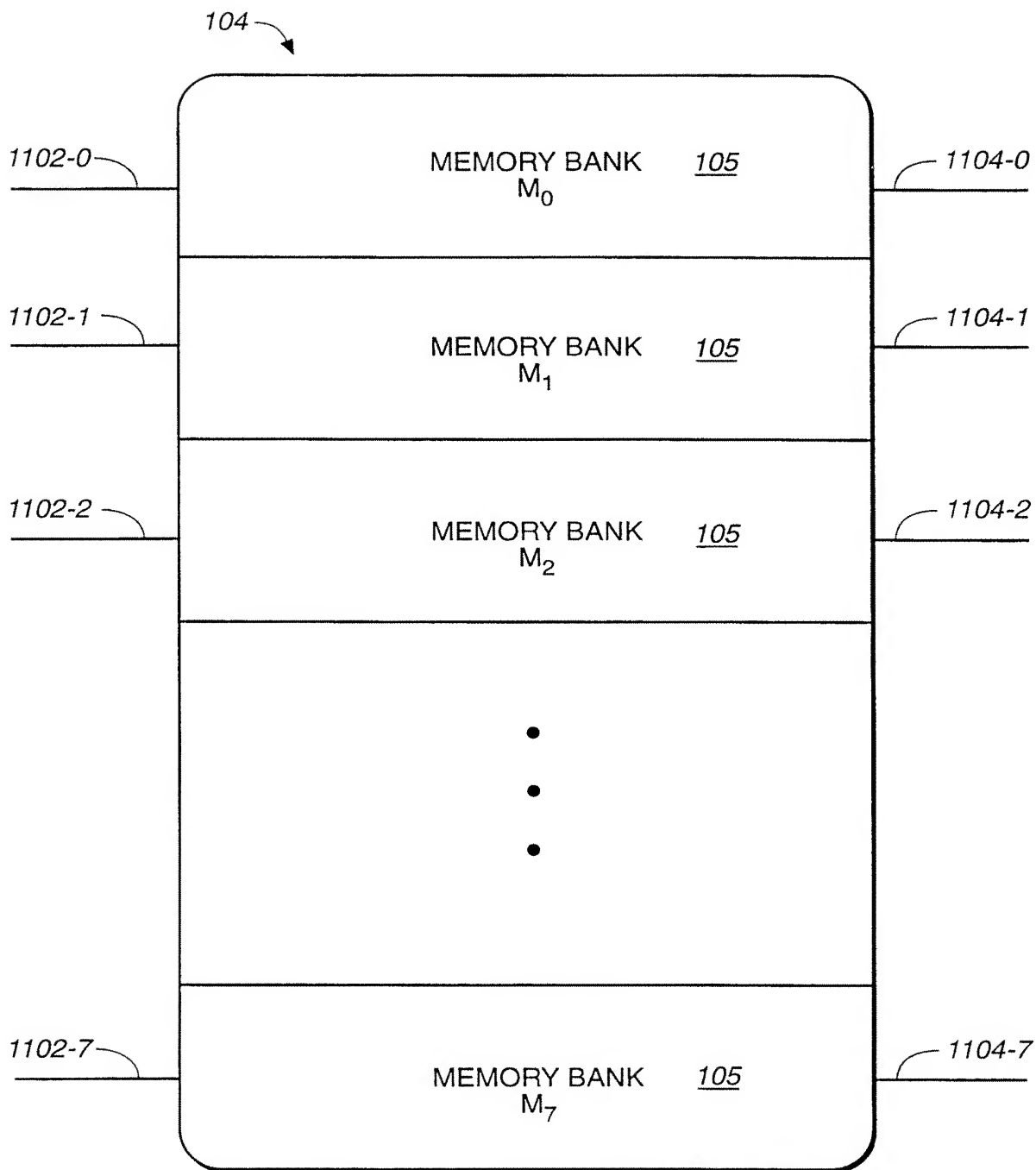
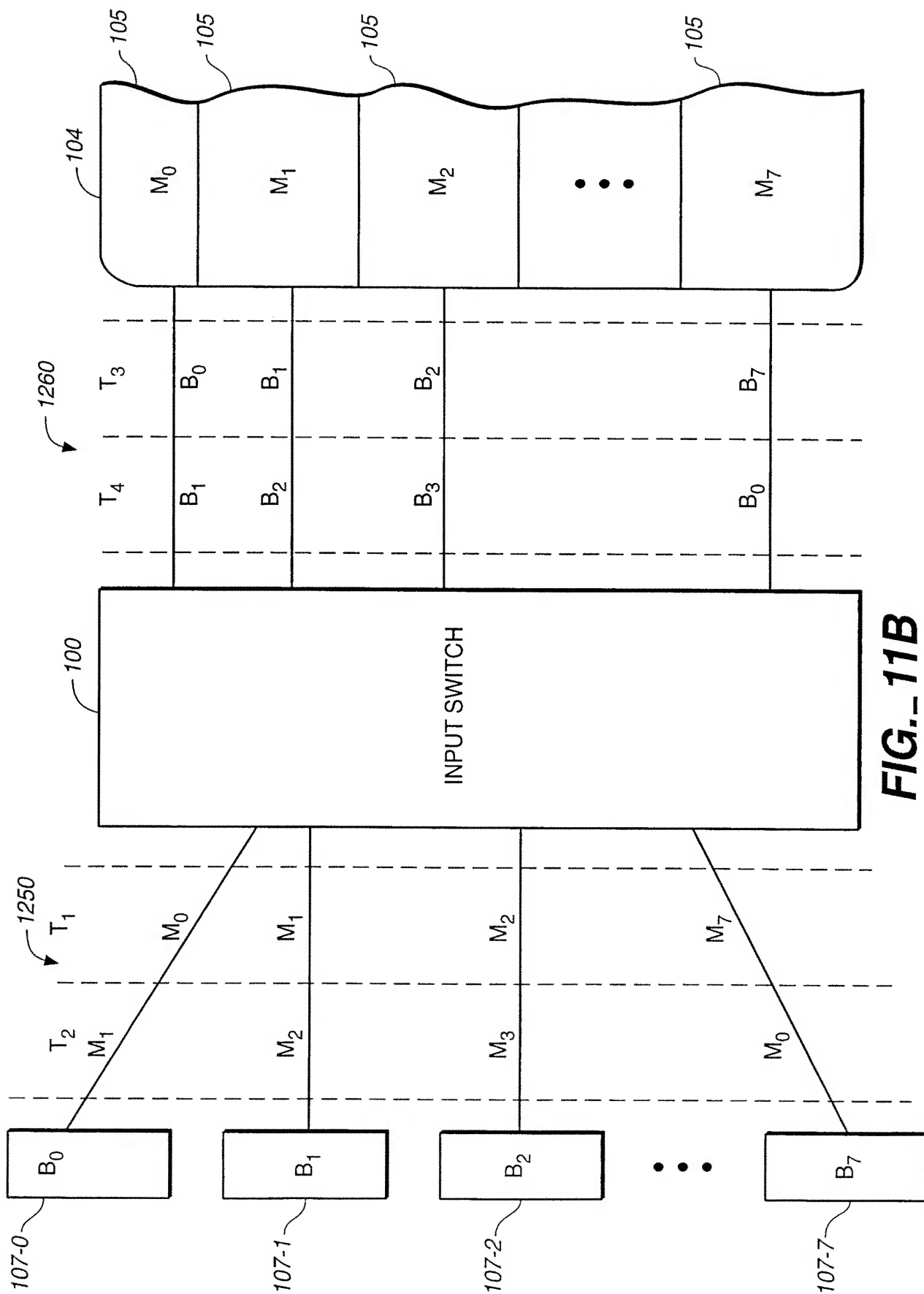


FIG. 11A



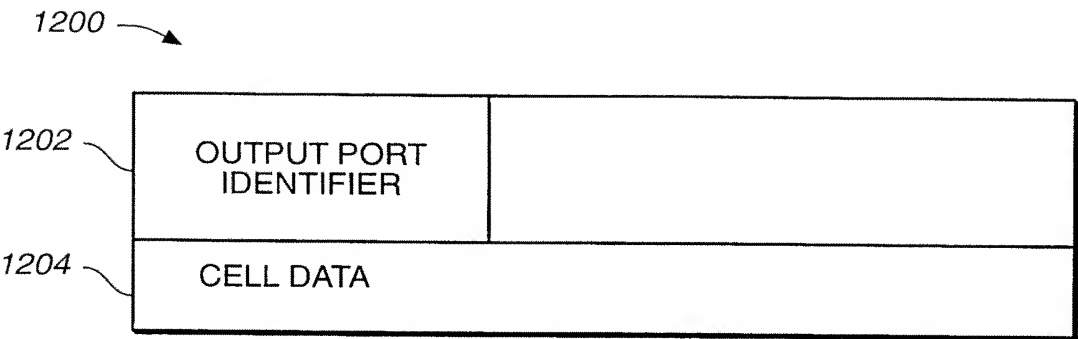


FIG._12

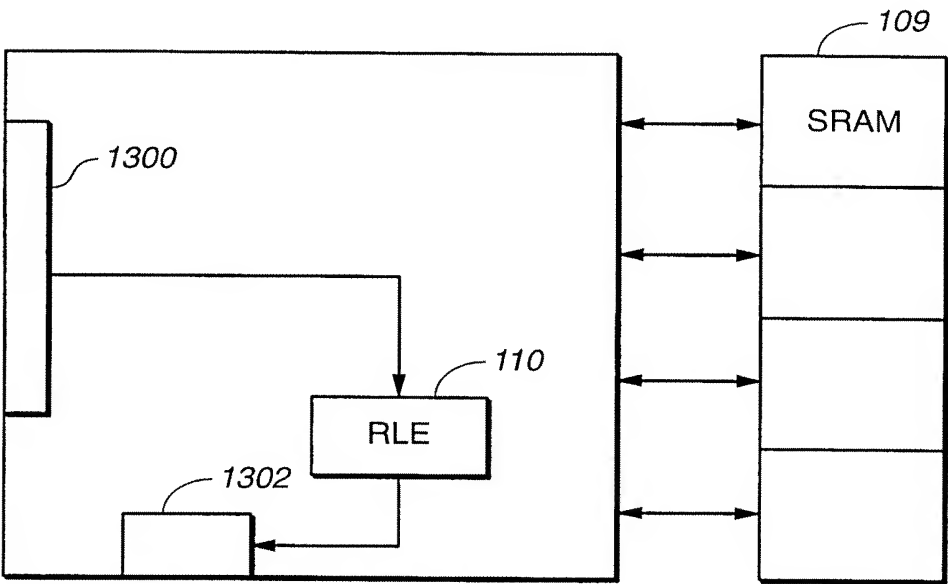


FIG._13

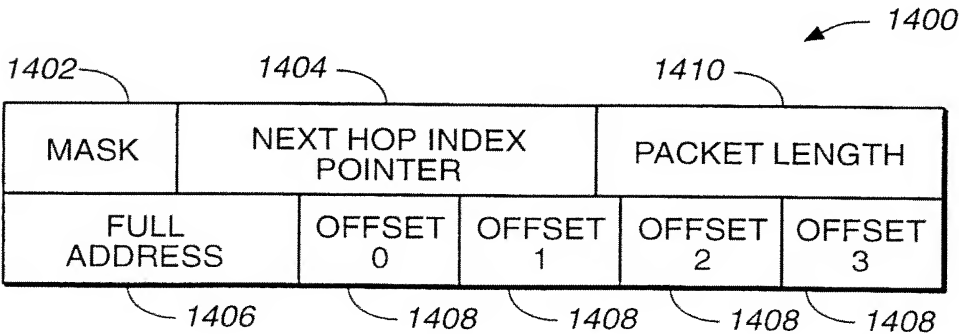


FIG._14

13 / 18

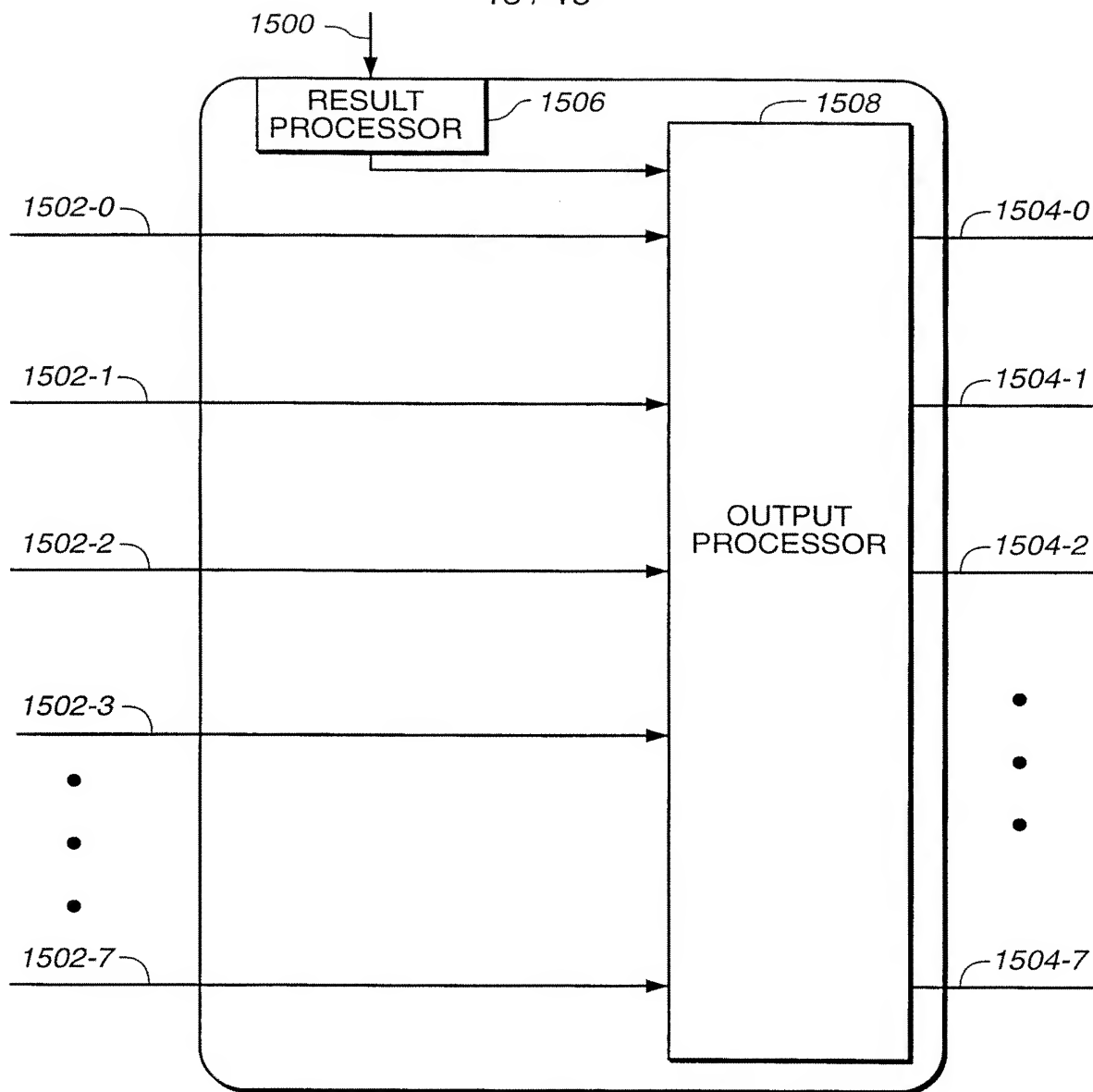


FIG. 15

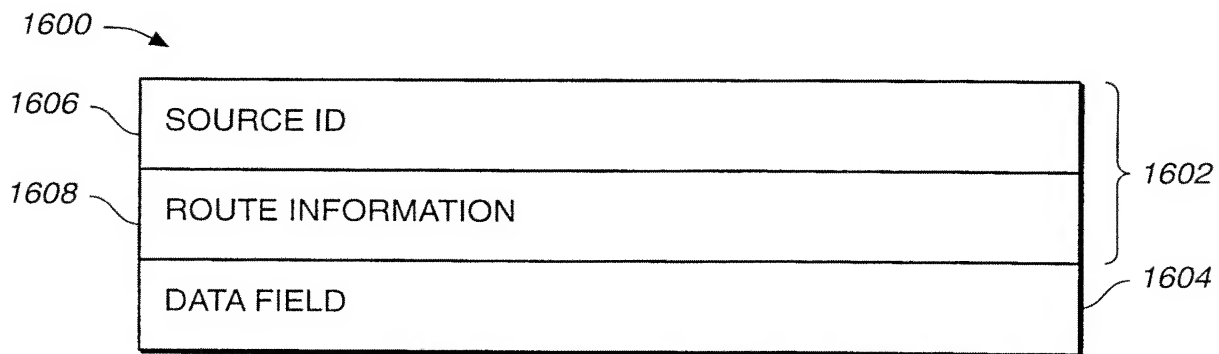
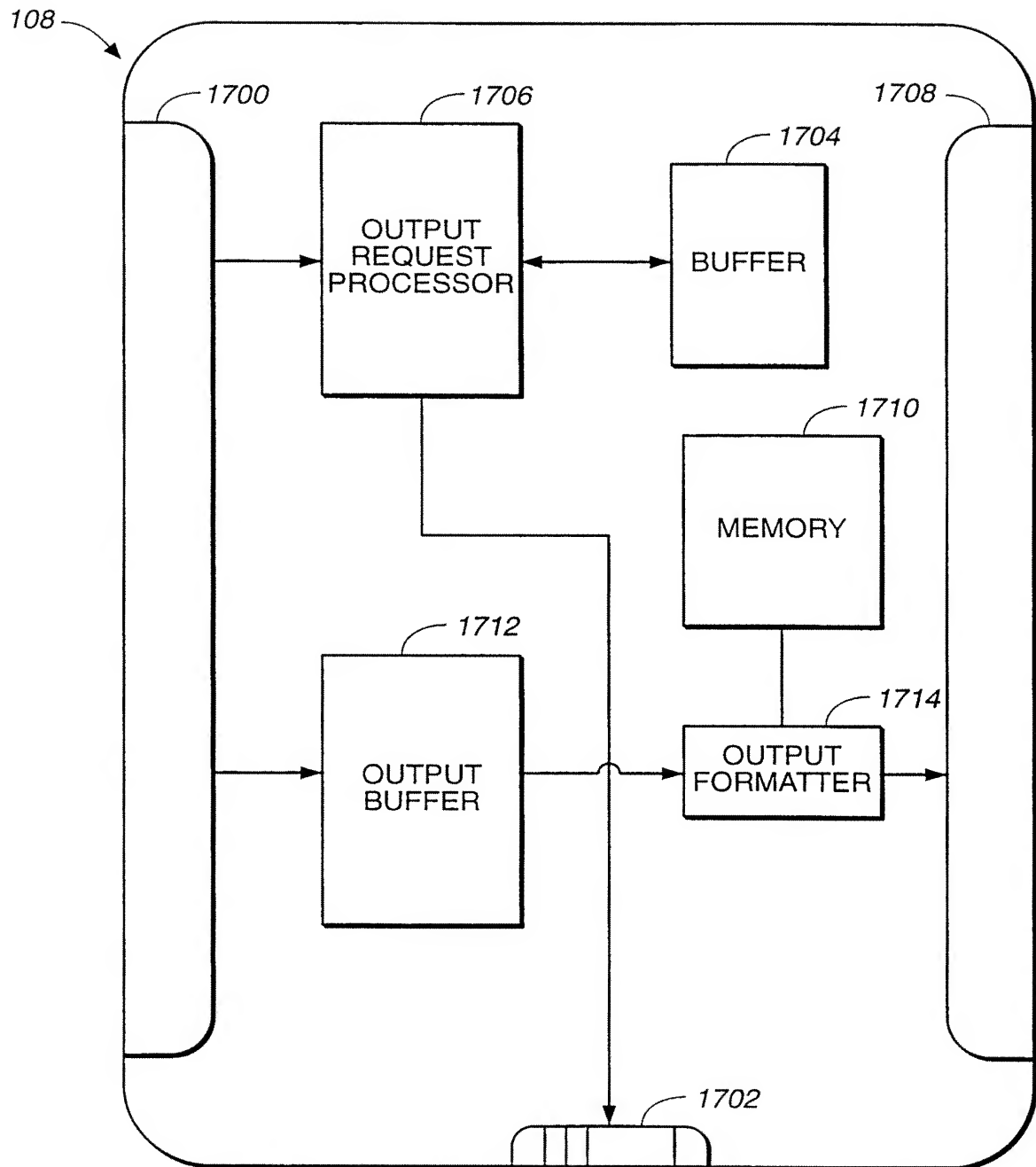
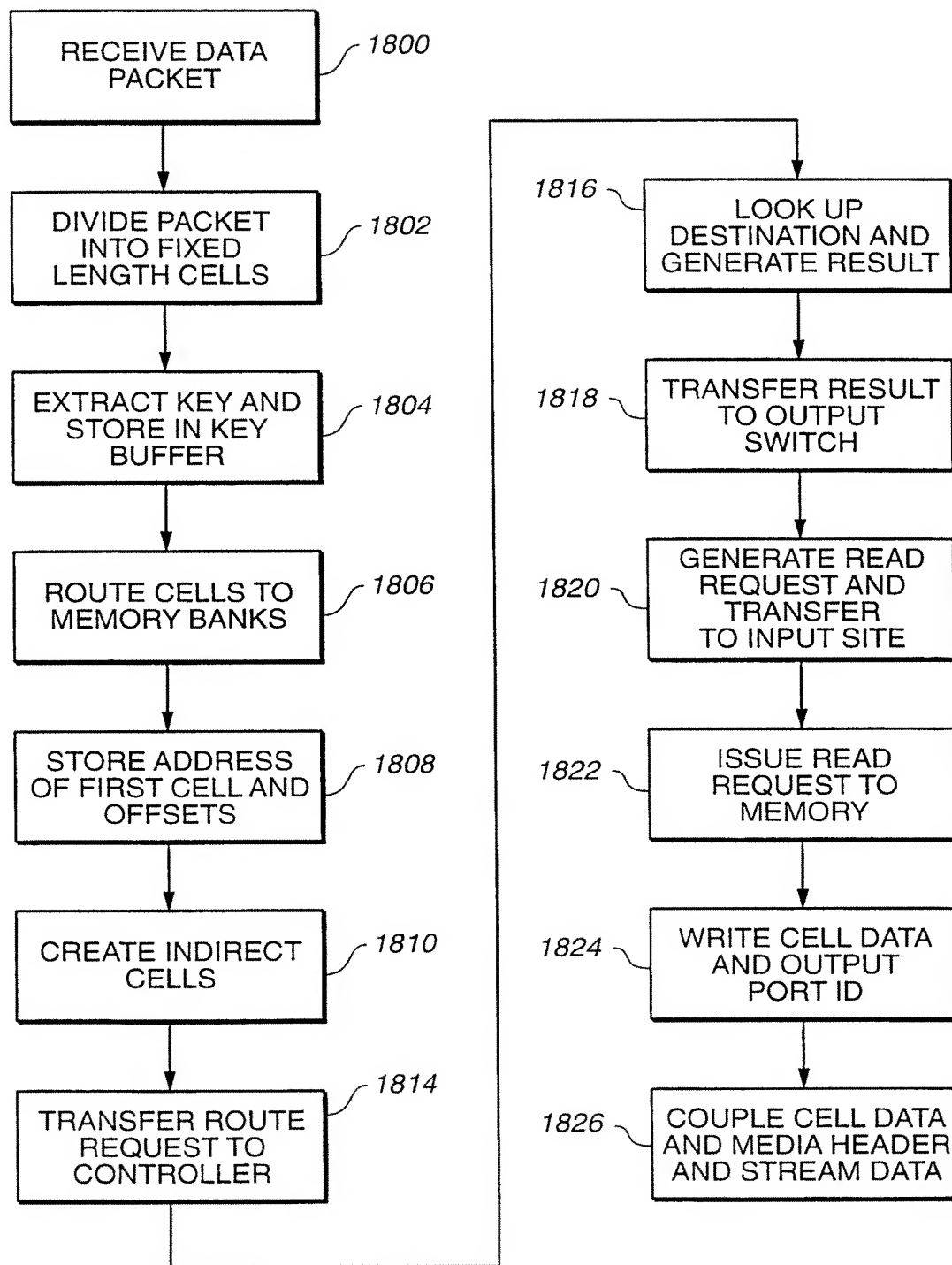


FIG. 16

**FIG. 17**

15 / 18

**FIG. 18**

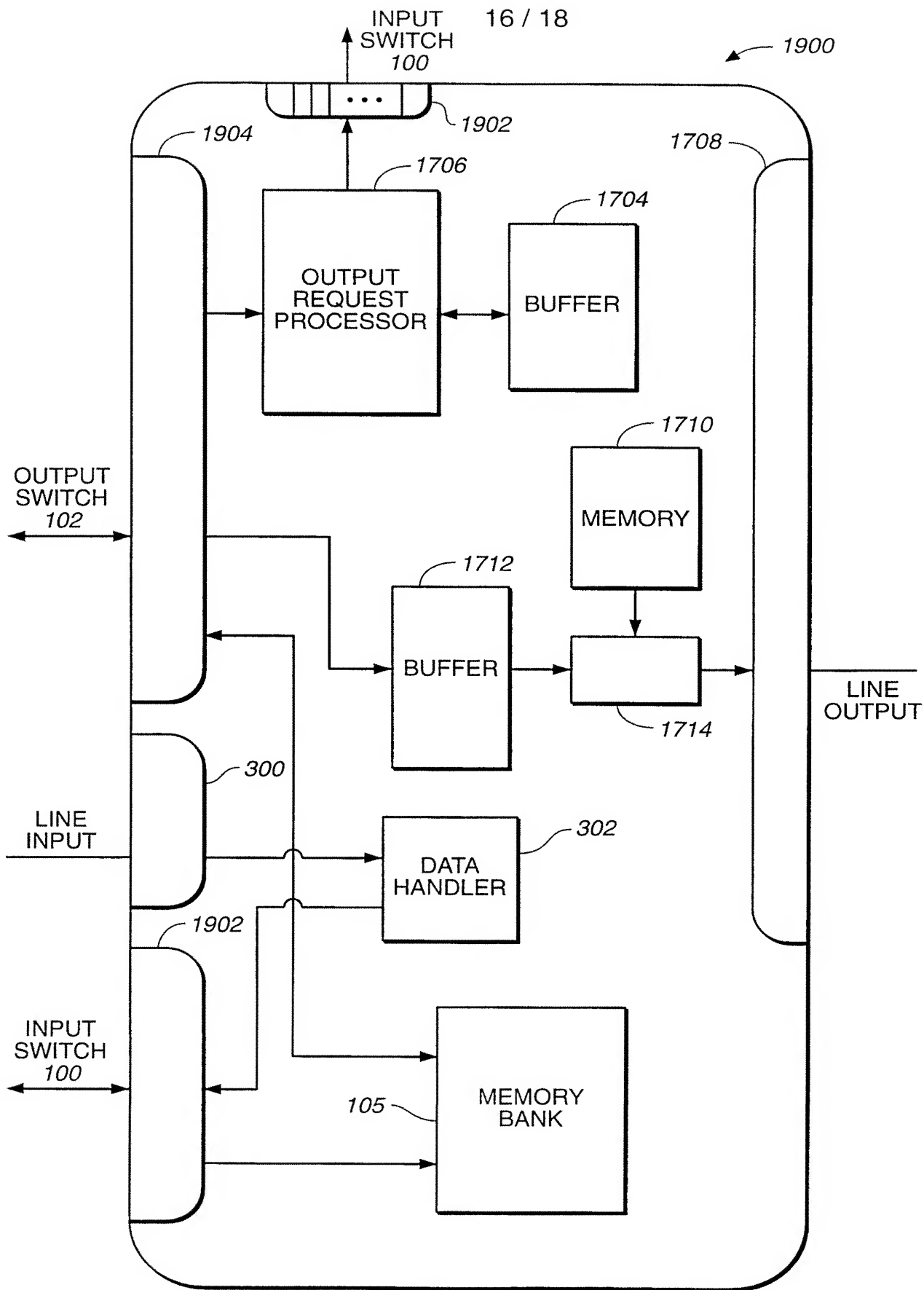
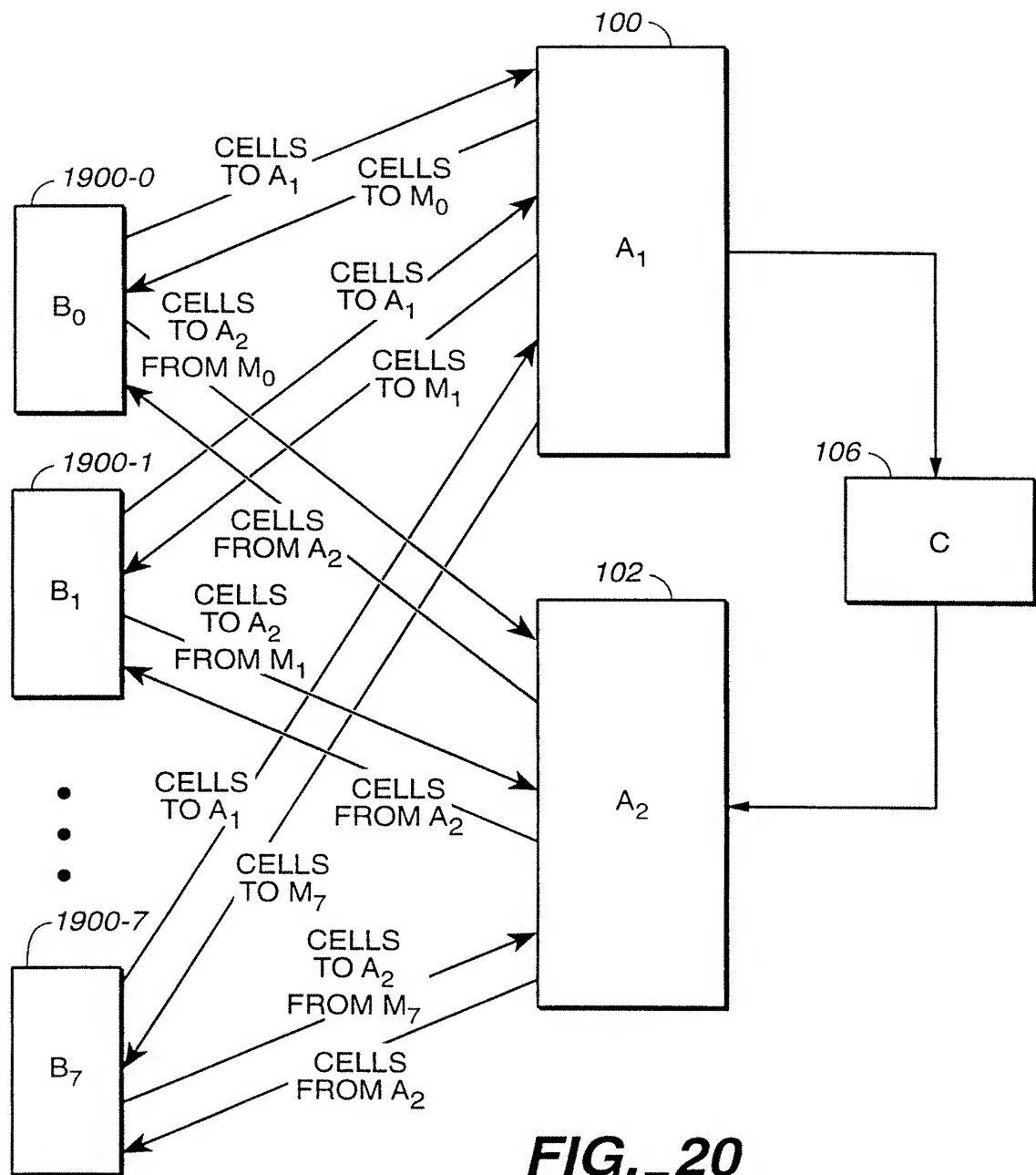


FIG. 19



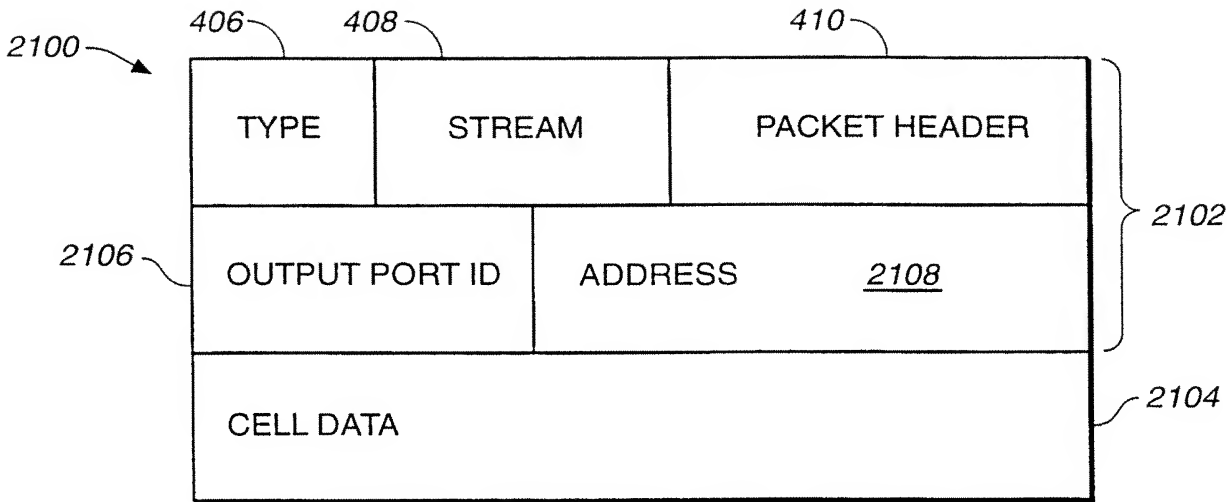


FIG._21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/23285

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :H04L 12/56

US CL :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 370/389,401,402,407,412,413,414,415,416,417,428,429; 395/280,844,873,917,909; 364/238,240,221

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
NONEElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)
NONE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ----- Y	US 5,491,694 A (OLIVER ET AL.) 13 February 1996, see columns 1-6 and figures 1 and 3.	1,4-12,14,17-25 ----- 2-3,13,15-16
Y	US 5,448,702 A (GARCIA ET AL) 05 September 1995, see column 6 lines 20-68.	2-3,13,15-16



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
B earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

09 APRIL 1998

Date of mailing of the international search report

28 JUL 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

DANG TON

Telephone No. (703) 305-4739

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/23285

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

370/389,401,402,407,412,413,414,415,416,417,428,429; 395/280,844,873,917,909; 364/238,240,221

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号
特表2000-516423
(P2000-516423A)

(43) 公表日 平成12年12月5日 (2000.12.5)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード (参考)
H 0 4 L 12/56		H 0 4 L 11/20	1 0 2 D
12/28			D

審査請求 有 予備審査請求 有 (全 52 頁)

(21) 出願番号 特願平10-527945
(86) (22) 出願日 平成9年12月16日 (1997.12.16)
(85) 翻訳文提出日 平成11年6月16日 (1999.6.16)
(86) 国際出願番号 PCT/US 97/23285
(87) 国際公開番号 WO 98/27660
(87) 国際公開日 平成10年6月25日 (1998.6.25)
(31) 優先権主張番号 08/767, 576
(32) 優先日 平成8年12月16日 (1996.12.16)
(33) 優先権主張国 米国 (US)
(31) 優先権主張番号 08/844, 171
(32) 優先日 平成9年4月18日 (1997.4.18)
(33) 優先権主張国 米国 (US)

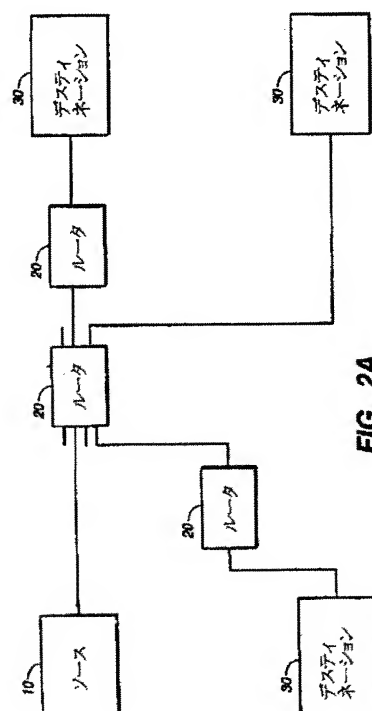
(71) 出願人 ジュニバー ネットワークス
アメリカ合衆国, カリフォルニア 95051,
サンタ クララ, ジェイ ストリート
3260
(72) 発明者 シンドウ, プラディーブ エス.
アメリカ合衆国, カリフォルニア 94040,
マウンテン ビュー, モンタルト ドライ
ブ 1557
(72) 発明者 アナンド, ラマリンガム ケイ.
アメリカ合衆国, カリフォルニア 95135,
サン ノゼ, トスカーナ コート 3096
(74) 代理人 弁理士 小橋 一男 (外1名)

最終頁に続く

(54) 【発明の名称】 高速スイッチング装置

(57) 【要約】

ネットワークにおいてソース (10) とデスティネーション (30) との間でデータパケットをスイッチングするルーター (20) が複数個の入力ポートを有しており、各入力ポートはデータハンドラーを有している。該データハンドラーは、データパケットを1個又はそれ以上の固定長セルに分割する。ルーター (20) は、少なくともそのうちの1つがデータパケットをデスティネーション (30) に対して経路付けするためのものである複数個の出力ポートと、複数個のメモリバンクに分割されているメモリとを有している。入力スイッチは入力ポートから固定長セルを受取り且つ単一の出力スイッチが前記メモリから受取ったセルを適宜の出力ポートに対して経路付けする。



【特許請求の範囲】

1. ネットワークにおけるソースとデスティネーションとの間でデータパケットをスイッチングするルーターにおいて、

データハンドラーを具備する入力ポートであって、前記ソースからのデータパケットを受取り前記データハンドラーが該データパケットを1個又はそれ以上の固定長セルへ分割する入力ポート、

前記データパケットを前記デスティネーションへルーチングする出力ポート、
複数のメモリバンクへ分割されているメモリ、

前記入力ポートから前記固定長セルを受取り且つセルスロットタイムスパンにおける単一のセルを各メモリバンクへルーチングする入力スイッチ、

前記メモリから受取ったセルを前記出力ポートに対してルーチングする出力スイッチ、
を有するルーター。

2. 請求項1において、前記入力スイッチが前記メモリ内の隣接していない位置から前記データパケットの検索を可能とするために前記データパケット内のセルをリンクさせるリンクエンジンを有していることを特徴とするルーター。

3. 請求項2において、更に、1個又はそれ以上の間接セルを発生する間接セル発生器を有してお

り、前記リンクエンジンが前記データパケットの連続するセルが格納されている前記メモリ内の位置をトラッキングし且つ間接セル内に格納するために前記データパケット内の各セルのメモリ内のアドレスを供給するルーター。

4. 請求項1において、前記入力スイッチが、入力ポートからの連続するセルが前記メモリ内の連続するバンクへ書込まれるように、前記メモリに対するデータパケットの書込を時分割多重処理するルーター。

5. 請求項1において、前記入力スイッチが、前記データパケットに関連する入力スイッチにおいて受取られた最初のセルからキー情報を抽出するキー読取エンジンを有しており、前記ルーターは、更に、前記入力スイッチへ結合されており且つそれから前記キー情報を受取る制御器を有しており、前記制御器は前記

前記ソースからデータパケットを受取るための入力ポートであって、前記データパケットを固定長セルへ分割するデータハンドラーを具備している入力ポート、

複数のメモリバンクに分割されているメモリ、

リンクエンジンを具備する入力スイッチであって、セルスロットタイムスパンにおいて前記入力ポートから高々単一セルを受取り且つ前記入力ポートからセルスロット内の前記メモリへ高々単一のセルをルーチングし、前記入力ポートからの連続するセルが前記メモリ内の連続するバンクへ書込まれるように前記メモリに対するセルの書込を時分割多重化処理し、前記データパケット内のセルをリンクさせるリンクエンジンが前記メモリ内の隣接していない位置からデータパケットの検索をすることを可能とする入力スイッチ、

前記データパケットと関連するデスティネーション情報をデコードし、本ルーターを介しての前記データパケットのルーチングを画定する通知を出力する制御器、前記制御器から前記通知を受取り且つメモ

リから前記出力ポートへ前記データパケットの転送を開始させる結果プロセサを具備している出力ポート、

前記メモリから受取ったセルを前記出力ポートに対してルーチングする出力スイッチ、
を有しているルーター。

14. ネットワークにおけるソースとデスティネーションとの間でデータパケットをスイッチングするルーターにおいて、

各々がデータハンドラーを具備している複数の入力ポートであって、第一入力ポートが前記ソースからのデータパケットを受取り、前記第一入力ポートのデータハンドラーが前記データパケットを1個又はそれ以上の固定長セルへ分割する複数の入力ポート、

少なくともそのうちの1つが前記データパケットを前記デスティネーションへルーチングさせるものである複数の出力ポート、

複数のメモリバンクへ分割されているメモリ、

入力スイッチから受取ったキー情報からデスティネーション情報をデコードし且つ前記メモリから前記出力ポートへ前記データパケットのルーチングを画定する通知を出力するルーター。

6. 請求項5において、前記出力ポートは、前記制御器からの通知を受取り且つ前記メモリから前記出力ポートへのデータパケットの転送を開始させる結果プロセサを有しているルーター。

7. 請求項1において、前記入力スイッチが、前記メモリから前記出力スイッチへの転送をスケジューリングする予約テーブルを有しているルーター。

8. 請求項7において、更に、前記入力スイッチ及び出力スイッチへ結合されており前記入力スイッチから受取られたデスティネーション情報をデコードし且つ前記メモリから前記出力ポートへのデータパケットのルーチングを画定する通知用前記出力スイッチへ出力する制御器を有しているルーター。

9. 請求項8において、前記出力スイッチが前記通知を前記出力ポートへ経路付けし且つその後前記出力ポートが前記出力スイッチを介してメモリから前記出力ポートへデータパケットを転送するために前記入力スイッチに対して要求を発行するルーター。

10. 請求項9において、前記出力ポートからの要求が前記予約テーブル内に格納されるルーター。

11. 請求項10において、メモリから前記出力スイッチへセルを転送するための要求が時分割多重化処理され、従って1つのスロットタイムスパン期間中に、高々単一の読取要求が処理するために前記メモリ内の各バンクに対して発行されるルーター。

12. 請求項9において、前記メモリが1つのセルスロットタイムスパンにおいてバンク当たり高々

単一のセルを出力するルーター。

13. ネットワークにおいてソースとデスティネーションとの間でデータパケットをスイッチングするルーターにおいて、

1つ又はそれ以上の入力ポートから固定長セルを受取り且つセルスロットタイムスパンにおいて各メモリバンクに対し単一のセルをルーチングさせる入力スイッチ、

前記メモリから受取ったセルを適宜の出力ポート

へルーチングさせる出力スイッチ、

を有しているルーター。

15. 請求項14において、前記入力スイッチが、前記メモリ内の隣接していない位置からのデータパケットからの検索を可能とさせるために前記データパケット内のセルをリンクさせるリンクエンジンを有しているルーター。

16. 請求項15において、更に、1個又はそれ以上の間接セルを発生する間接セル発生器を有しており、前記リンクエンジンが、前記データパケットの連続するセルが格納されている前記メモリ内の位置をトラッキングし且つ間接セル内に格納するために前記データパケット内の各セルのメモリ内のアドレスを供給するルーター。

17. 請求項14において、前記入力スイッチが、任意の入力ポートからの連続するセルが前記メモリ内の異なるバンクへ書込まれるように、前記メモリに対するセルの書込を時分割多重化処理するルーター。

18. 請求項14において、前記入力スイッチが前記データパケットに関連する入力スイッチにおいて受取られた最初のセルからキー情報を抽出するためのキー読取エンジンを有しており、本ルーターは、更に、前記入力スイッチへ結合されており且つ

それからキー情報を受取る制御器を有しており、前記制御器は前記入力スイッチから受取ったキー情報からデスティネーション情報をデコードし且つ前記メモリから適宜の出力ポートへの前記データパケットのルーチングを画定する通知を出力するルーター。

19. 請求項18において、前記出力ポートは、前記制御器から前記通知を受取り且つ前記メモリから適宜の出力ポートへのデータパケットの転送を開始さ

せる結果プロセサを有しているルーター。

20. 請求項14において、前記入力スイッチが前記メモリから前記出力スイッチへの転送をスケジュールする予約テーブルを有しているルーター。

21. 請求項20において、更に、前記入力スイッチと出力スイッチとに結合されており前記入力スイッチから受取ったデスティネーション情報をデコードし且つ前記メモリから適宜の出力ポートへデータパケットのルーチングを画定する通知を前記出力スイッチへ出力する制御器を有しているルーター。

22. 請求項21において、前記出力スイッチが前記通知を前記出力ポートに対して経路付けし且つその後前記出力ポートが前記出力スイッチを介してメモリから適宜の出力ポートへデータパケットを転送するために前記入力スイッチに対して要求を発行するルーター。

23. 請求項22において、前記出力ポートからの要求が前記予約テーブル内に格納されるルーター。

24. 請求項23において、メモリから出力スイッチへセルを転送するための要求が時分割多重化処理され、従って1つのセルスロットタイムスパン期間中に、高々単一の読取要求が処理するために前記メモリ内の各バンクへ発行されるルーター。

25. 請求項22において、前記メモリが1つのセルスロットタイムスパンにおいて、バンク当たり高々単一のセルを出力するルーター。

能な条件のことを意味する。各入力バッファ4の寸法は、例えば、ライン入力速度、ルックアップ処理の速度、スイッチング装置に対するブロッキング特性等の多数のファクタに依存する。

然しながら、これらのタイプのルーターは多くの観点において効率が悪い。各入力ポートは専用の入

力バッファを有しており、且つ入力ポート間のメモリ共有はその設計において設けられていない。各入力バッファは、与えられたポートに対する最大の処理能力条件を満足するために寸法決定されねばならない。然しながら、設計トレード（コスト）が、しばしば、各ポートに対してより小型のバッファとすることを必要とする。より小型のバッファの場合には、ブロッキング条件のためにパケットがドロップ即ち落されて削除される可能性が発生する。ルーターにおいては典型的に余分のメモリ能力が存在しているが（入力ポートの様々な利用に起因して）、その余分のものを利用する手段が与えられていない。

パケットのドロップ即ち落されることを防止するために、設計者は「非ブロッキング」ルーターを開発した。図1bを参照すると、従来の「非ブロッキング」ルーターは、各々が入力バッファ（メモリ）4を包含している複数個の入力ポート2と、スイッチング装置6と、各々が出力バッファ9を具備する複数個の出力ポートとを有している。ブロッキング条件を開示するために、各出力ポート8は出力バッファ9を包含する形態とされている。各出力ポートは同時にパケットを出力すると共に後の時間において出力するための新たなパケットを受取ることが可能である。出力バッファの寸法が充分に大きなも

のである場合には、データパケットがドロップ即ち落されることはない。

然しながら、これらの設計はメモリ容量及びコストの点において更に効率が悪い。この場合も、各出力ポートは専用の出力バッファを有しており且つ出力ポート間のメモリ共有はこの設計においては設けられていない。各出力バッファは、与えられたポートに対する最大の処理能力条件を満足するために（その「非ブロッキング」特性を維持するために）寸法設定されねばならない。典型的に、ルー

【発明の詳細な説明】

高速スイッチング装置

発明の詳細な説明

背景

本発明は、大略、データルーチングシステムに関するものであって、更に詳細には、ネットワークを介してパケットのルーチング即ち経路付けを行う方法及び装置に関するものである。

パケット交換通信システムにおいては、ルーターが1つのポート上でデータ又は制御情報を包含するパケットを受取り、且つ該パケット内に含まれているデスティネーション情報に基づいて、該パケットを別のポートからデスティネーション（又は中間デスティネーション）へ経路付けを行うスイッチング装置である。

従来のルーターは、特定のパケットに対して適切な出力ポートを決定するために、パケット内の最初のデータブロック内に含まれているヘッダー情報を評価することによってこのスイッチング機能を実行する。

ルーターを介してのパケットの効率的なスイッチングが極めて重要である。図1aを参照すると、従来のルーターは、各々が入力バッファ（メモリ）4を包含している複数個の入力ポート2と、スイッチ

ング装置6と、複数個の出力ポート8とを有している。

入力ポート2において受取られたデータパケットは、少なくとも一時的に入力バッファ4内に格納され、一方各パケットと関連するデスティネーション（宛先）情報がデコードされてスイッチング装置6を介しての適切なスイッチングを決定する。明らかに、デコードプロセッサがパケットが受取られる速度と比較して長くかかり過ぎる場合には、より大きなメモリ要素が必要とされる。更に、パケットは、スイッチング装置が接続を行うことが不可能である場合には、デスティネーション情報がデコードされた後に入力バッファ内に留まることが強制される場合がある。ブロッキング（blocking）は、所望の出力ポートが使用不可能（ポートがビジーであり、例えば、異なる入力ポートからの別のパケットをルーチングしている）であるために、スイッチにおいて接続を形成することが不可

ターにおいては更に余分のメモリ能力が存在しているが（入力ポート及び出力ポートの多様な利用に起因して）、その余分なものを利用する手段が与えられていない。これらのタイプの装置を介して移動されるデータの量をサポートするために必要とされるものよりも2倍の量及び帯域幅のメモリが使用されねばならない。

発明の要約

一般的に、1つの側面においては、本発明は各々がデータハンドラーを包含する複数個の入力ポートを具備するネットワークにおいてソース（発信元）とデスティネーション（宛先）との間でデータパケットをスイッチング即ち交換するためのルーターを提供している。該データハンドラーはデータパケットを1個又はそれ以上の固定長のセルへ分割する。

本ルーターは複数個の出力ポートを有しており、少なくともそのうちの1つはデータパケットをデスティネーションに対してルーチング即ち経路付けするためのものであり且つ本ルーターは複数個のメモリバンクへ分割されているメモリを有している。入力スイッチが入力ポートから固定長のセルを受取り且つセルスロットタイムスパン内の単一のセルを各メモリバンクへ書込む。出力スイッチがメモリから受取ったセルを適宜の出力ポートに対して経路付けする。

本発明の側面は以下の特徴を包含している。入力スイッチはメモリ内の非隣接位置からデータパケットの検索を可能とするためにデータパケット内のセルをリンクするためのリンクエンジンを有している。本ルーターは、更に、1個又はそれ以上の間接セルを発生するための間接セル発生器を有している。該リンクエンジンは、データパケットの連続するセルが格納されているメモリ内の位置をトラッキング即ち追いつき、且つ間接セル内に格納するためにデータパケット内の各セルのメモリ内のアドレスを供給する。

本入力スイッチは、メモリに対するデータパケットの書込を時分割多重し、従って入力ポートからの連続したセルはメモリ内の連続したバンクへ書込ま

れる。本入力スイッチは、データパケットと関連する入力スイッチにおいて受取

られた最初のセルからのキー情報を抽出するためのキー読取エンジンを有している。本ルーターは、更に、それからキー情報を受取るために入力スイッチへ結合されている制御器を有している。該制御器は入力スイッチから受取られたキー情報からデスティネーション情報をデコードし且つメモリから出力ポートへデータパケットのルーチング即ち経路付けを画定する通知を出力する。

出力ポートは制御器からの通知を受取るための結果プロセッサを有しており且つメモリから出力ポートへのデータパケットの転送を開始させる。本入力スイッチはメモリから出力スイッチへの転送をスケジュールするための予約テーブルを有している。該出力スイッチは、出力ポートへの該通知の経路付けを行い、その後、出力ポートは入力スイッチに対してメモリからのデータパケットを出力スイッチを介して出力ポートへ転送することの要求を発行する。出力ポートからの要求は予約テーブル内に格納される。メモリから出力スイッチへセルを転送するための要求は時間ドメイン多重化され、従って1つのセルスロットタイムスパン期間中に、高々、単一の読取要求が処理するためにメモリ内の各バンクに対して発行される。メモリは、1つのセルスロットタイ

ムスパン内においてバンク当たり高々単一のセルを出力する。

本発明の1つの利点は、スイッチを介してのバケットを効率的に管理し且つ経路付けするスイッチングアーキテクチャを設けることによって高価な高速メモリ内にバケットを格納することの必要性なしに、ライン速度においてルーターを介してバケットをスイッチ即ち交換させることが可能であるということである。

その他の利点及び特徴は以下の説明及び請求の範囲から明らかとなる。

図面の簡単な説明

図1 a及び1 bは従来のルーター装置のブロック図である。

図2 aは本発明の1実施例に基づくデータルーチングシステムの概略ブロック図である。

図2 bは本発明の1実施例に基づくルーターの概略ブロック図である。

図3は本発明の1実施例に基づく入力ポートの概略ブロック図である。

図4 aは本発明の1実施例に基づく入力ポートから入力スイッチへ転送するタ

。

図18は本発明の1実施例に基づくルーターを介してのバケットのルーチングのプロセスに対するフローチャートである。

図19は本発明の1実施例に基づくルーターの概略ブロック図であって、入力ポートからその適切な出力ポートへバケットを経路付けする場合にルーターによって実行される動作の流れを追従するためのシーケンス流れ情報を包含する概略ブロック図で

ある。

図20は本発明の1実施例に基づく多機能ポートの概略ブロック図である。

図21は本発明の1実施例に基づく多機能ポートを包含するルーターの概略ブロック図である。

詳細な説明

図2 aを参照すると、バケットスイッチングシステムにおいて、ソース（発信元）10が1つ又はそれ以上のデスティネーション（宛先）30へバケットを送信するために1個又はそれ以上のルーター20へ接続されている。各ルーターは、種々のソース及びデスティネーションへ接続されている複数のポートを有している。ソース10からのバケットは、そのデスティネーションに到達する前に、1個を超えるルーターを介して通過することが可能である。

図2 bを参照すると、各ルーター20は入力スイッチ100と、出力スイッチ102と、1個又はそれ以上のメモリバンク105を包含するメモリ104と、制御器106と、複数の入力及び出力ポート107及び108とを有している。ルーチングテーブルを格納するために、制御器メモリ109が制御器106と関連している。入力スイッチ100は各入力ポート107へ接続されており、一方出力スイッチ102はルーター20における各出力ポート1

08へ接続されている。1実施例においては、ルーター20は8個の入力及び出力ポート107及び108を有している。

動作について説明すると、バケットが入力ポート107において受取られ、入

力ミング及び順番を包含するルーターの概略ブロック図である。

図4 bは本発明の1実施例に基づく入力ポートと

入力スイッチとの間で転送されるセルに対するデータ構造である。

図5 aは本発明の1実施例に基づく入力スイッチの概略ブロック図である。

図5 bは本発明の1実施例に基づく入力ポートからメモリへ転送するタイミング及び順番を包含するルーターの概略ブロック図である。

図6は本発明の1実施例に基づくキーバッファ内に格納されているルート要求に対するデータ構造である。

図7は本発明の1実施例に基づく間接セルに対するデータ構造である。

図8は本発明の1実施例に基づく入力スイッチとメモリバンクとの間で転送されるセルに対するデータ構造である。

図9は本発明の1実施例に基づく予約テーブルの概略ブロック図である。

図10は本発明の1実施例に基づく予約テーブルをローディングするプロセスのフローチャートである。

図11 aは本発明の1実施例に基づくメモリの概略ブロック図である。

図11 bは本発明の1実施例に基づく入力ポートからメモリへの転送のタイミング及び順番を包含す

るルーターの概略ブロック図である。

図12は本発明の1実施例に基づくメモリバンクから出力スイッチへのセル出力に対するデータ構造である。

図13は本発明の1実施例に基づく制御器の概略ブロック図である。

図14は本発明の1実施例に基づく制御器から出力スイッチへ転送される出力要求に対するデータ構造である。

図15は本発明の1実施例に基づく出力スイッチの概略ブロック図である。

図16は本発明の1実施例に基づく出力スイッチから出力ポートへ転送されるセルに対するデータ構造である。

図17は本発明の1実施例に基づく出力ポートに対する概略ブロック図である

カスイッチ100へ転送され、且つ一時的にメモリ104内に格納される。該バケットがスイッチ100によって受取られると、そのバケット内の最初のデータブロックからキーが読取られ且つ制御器106へ転送される。そのキーは、バケット内のデータの最初のブロックと関連するヘッダーフィールドから派生されるデスティネーション情報及びその他の情報（ソースID、流れID等）を包含している。

制御器106内のルートルックアップエンジン110は、キー情報にも基づいてツリーをベースとしたサーチを実行し且つデスティネーションと関連している出力ポートを含む結果を返す。その結果はルーター20を介してバケットを経路付けするためのその他の情報（ソースID、流れID、バケット長等）と結合され且つ制御器106から出力スイッチ102に対する通知として供給される。出力スイッチ102はその通知を識別された出力ポート108へ転送する。該通知情報を受取ると、出力ポート108はメモリ104から出力スイッチ102を介して適切な

出力ポート108へバケットの転送を開始する。

図3を参照すると、各入力ポート107はライン入力インターフェース300と、データハンドラー304と、セル出力ポート306とを有している。バケットはライン入力インターフェース300において受取られる。バケットが受取られると、データハンドラー302は受取ったバケットを固定長セルへ分割する。本発明の1実施例においては、各セルの長さは80バイトであり16バイトが内部ヘッダー（制御情報）であり且つ64バイトがセルデータである。データハンドラーが入って来るバケットを固定長セルへ分割すると、それはセル出力ポート306を介して同期的にセルを入力スイッチ100へ出力する。

次に、図4 aを参照すると、単一のセル450が入力ポート107から各セルスロット「T」において入力スイッチ100へ転送される。与えられたセルスロット「T」に対して、入力ポート107は全部で「N」のセルを受取り、尚「N」は入力ポートの数に等しい。

入力ポート107から入力スイッチ100へ転送される各セル400に対する

を介して受取った入力（セル）をサービス即ち処理する。即ち、与えられたセルスロットに対して、各入力ポートからの１個のセルがラウンドロビンデータハンドラ－５００において受取られ且つ、その後、メモリ１０４内のメモリバンク１０５へ次のセルスロットにおいて転送するために出力プロセサ５０５へ転送される。次のタイムセルスロットにおいて、データハンドラ－５００は、同一の入力ポートから受取った次のセルを異なるメモリバンクに対して転送するために出力プロセサ５０５へ転送する。１実施例においては、受取った次のセルはメモリアレイ内の次のメモリバンク（数値順において次）へ転送される。一方、同一の入力ポートから連続するセルの転送を制御するために別の時間依存性順列を使用することが可能である。

450-0乃至450-7が伝送線458の下方向へ転送され、尚各入力ポート107から入力スイッチ100へ1つである。セルスロットT3（時間においてセルスロットT4のすぐ前のもの）において、一連のセル452-0乃至452-7が伝送線458下方向へ転送され、尚各入力ポート107から入力スイッチ100へ1つである。

5 個を超えるデータセルが 1 個のバケット内に包含されている場合には、間接セルインジケータがセットされ、且つ最後のオフセットはそのバケットに関連する最初の間接セルが格納されているメモリ内のアドレスを表わす。間接セルは図 7 に関連して以下に更に詳細に説明する。バケットがメモリ内に格納された後に、キーバッファ 5 1 6 内の関連するエントリ（ルートルックアップ要求）が制御器インター

同一のバケットに対するセルのリンク形成又はスレッド形成は、上述したオフセット及び間接セルを使用して行われる。オフセットは、バケット内の最初の5個のセルをリンクするために使用され、一方間接セルはバケット内の残りのセルをリンクするために使用される。1実施例においては、1個のセルが5個以下のセルを包含している場合には、間接セルを使用することが必要とされることはない。間接セルプロセッサ510は与えられたバケットに対してメモリ内のセルのリンク形成を行う。間接セルプロセッサ510はメモリ104内に格納するために間接セルを発生する。間接セルはバケット内の隣接セルの間のメモリ空間における相対的なオフセットに関連するオフセット情報を包含している。間接セルプロセ

タイプフィールド406は入力ポートから転送されるべきセルのタイプを表わす。各セルスロット（1実施例においては20クロックサイクル）において、入力ポートがデータセル、間接セルブレースホルダー、又は遅延間接セルホルダーのいずれかを転送することが可能である。データセルは入って来るパケットと関連するデータを包含している。間接セルブレースホルダーは空のセルであり、且つメモリ104内のセルの格納に対する間接アドレス動作に関連して使用される。遅延間接セルブレースホルダーは、間接アドレス動作を必要とするデータストリームが、そのデータストリームと関連する最後の間接アドレス用セルをメモリ104へ書き込むための指定されている時間間の前時間において終了する場合に発生する。間接ブレースホルダー及び遅延間接ブレースホルダーの発生及び動作については図7に関連して後に更に詳細に説明する。

バケットヘッダーフィールド410は与えられたバケットに関連するヘッダー情報を包含しており且つスタートオフセット情報と、バケット長と、インターフェースインデックス情報とを有している。

入力スイッチ100内のラウンドロビンデータハンドラー500及び出力プロセッサ505は、伝送線460を介してセルをメモリ100へ転送する。セルスロットT2において理解することが可能であるように、出力プロセッサ505は1個のセル454-B乃至454-B'を単一セルスロットにおいて各メモリバンクへ出力する。「B」指定子はそれから特定のセルが受取られた入力ポートを表わす。各入力ポートからの1個のセルがセルスロット当たりメモリ104へ書込まれる。時間期間T1（セルスロットT2の前の1つのセルスロット）において、1個のセル456-B乃至456-B'が各メモリバンクへ書込まれる。ラウンドロビンデータハンドラー500は、同一の入力ポートからの連続したセルがメモリ104内の連続したメモリバンク105へ書込まれるように、出力プロセッサ505への転送を時分割多重処理する。

ラウンドロビンインターハンドラー 5 0 0 は、バケット内の最初のセルに関連するキー情報を決定するためのキー読取エンジン 5 1 4 と、同一のバケット内のセルをリンクするためのリンクエンジン 5 1 5 とを有している。

キー情報を読取るプロセスは当該技術において公知である。与えられたバケットに対してそのキーが決定された後に、それはバケット全体がメモリ104内に格納されるまで、入力スイッチ100内のキーバッファ516内に一時的に格納される。キーバッファ516内のエントリ600に対するデータ構造を図6に示してある。各エントリ600はキー602と、フルアドレス604と、オフセット606と、間接セルインジケータ608とを有している。

リンクエンジン 515 は、与えられたバケット内の最初のセルに対してメモリ内の開始アドレス（フルアドレス）を決定する。この開始アドレスはメモリ 104 内のバンク番号（ラウンドロビンデータハ

サは間接セルを形成する期間中に間接セルデータを格納するための間接セルメモリ 520 を有している。

次に、図 7 を参照すると、間接セル 700 に対するデータ構造は、リンク形成フィールド 702 と、複数のオフセットフィールド 704 と、最後のフ

ィールド 706 とを有している。リンク形成フィールド 702 は、セットされていない場合には、現在の間接セルが与えられたバケットに対する間接セルの遅延内における最後のセルであることを表わす。セットされている場合には、与えられたバケットに対して更なる間接セルが存在している。更なる間接セルが存在している場合には、最後のフィールド 606 はそのバケットと関連する次の間接セルのメモリ内の位置に対するオフセットを表わす。本発明の 1 実施例においては、各間接セルはメモリ内の 56 個のセルをリンクさせるための 56 個のオフセットデータブロックを有している。

上述したように、バケットが受取られると、リンクエンジンが最初の 5 個のセルを処理し且つ開始アドレスの形態におけるリンク情報及び 4 個のオフセットをキバプファ 516 内に格納する。5 個を超える数のセルが 1 個のバケット内に包含されている場合には、間接セルプロセサはリンクエンジンにとって代わり且つそのバケット内の残りのセルが格納されているメモリ内の位置に関連するオフセットを計算する。ラウンドロビンプロセサ 500 はメモリ 104 内の関連するメモリバンクへ転送するためにセルを出力プロセサ 505 へパスする。ラウンドロビンプロセサ 500 は処理中のバケットが 5 個を超

える数のセルを有している場合（最初のセル内に含まれているヘッダー情報に基づいて）、間接セルプロセサをイネーブル即ち動作可能状態とさせる。5 番目のセルをメモリへ書込む時に、間接セルプロセサ 510 はそれがバケット内の最後のセルであった場合に 5 番目のセルが書込まれるであろうメモリ内の位置と関連するアドレス（「間接セルアドレス」）を間接セルメモリ 520 内に格納する。間接セルアドレスは、それがフル即ち満杯である場合に（又はそのバケットの最後のセルが処理された場合に）間接セルが書込まれるべきメモリ内の位置を表わ

す。読取制御器 517 はメモリインターフェース 502 を出て入力スイッチ 100 からメモリ 104 内の個別的なメモリバンクへ流れる読取要求信号の転送を制御する。読取制御器 517 は出力ポートインターフェース 506 を介して各出力ポートからの読取要求を受取る。各要求のフォーマットは、ソース ID（出力ポート）及び読取られるべきメモリ内のフルアドレスを有している。各セルスロットにおいて、各出力ポートはメモリ 104 内のメモリ位置を読取るためにスイッチ 100 によって処理するための読取要求を発生することが可能であり、その結果メモリバンク 105（後のセルスロット上）から出力スイッチ 102 へのセルの読取（読取応答）が行われる。

入力スイッチ 100 から（出力プロセサ 505 を介して）メモリ 104 内のメモリバンク 105 へ転送されるセルのデータ構造を図 8 に示してある。各セルスロットにおいて、出力プロセサ 505 はセル 800 を発生し、それは読取要求ソースフィールド 802 と、読取アドレス 804 と、書込アドレス 806 と、データフィールド（入力ポート 107 から受取られたセルデータ）808 とを有している。読

取要求ソースフィールド 802 は、読取を要求する出力ポート（デスティネーション出力ポート 108）を表わす。出力プロセサ 505 は読取制御器 517 から読取要求を受取り且つその読取要求を同一のメモリバンクに対して意図されているラウンドロビンデータハンドラー 500 から受取られた何等かの書込要求とバンドル即ち結束させる。各セルスロットにおいて、出力プロセサ 505 はセル 800 を供給し、それはメモリ 104 内の各メモリバンク 105 に対する書込及び読取要求を有することが可能である。

読取制御器 517 は種々の出力ポート 108 からバケットを転送するための要求が受取られると予約テーブル 508 をロードする。予約テーブルは、全てのセルスロットにおいて、メモリ 105 の各バンクに対して単一の読取要求が発生されるようにロードされる。次に図 9 を参照すると、予約テーブル 508 は、メモリ 104 内の各メモリバンク 105 に対して 1 つづつ複数の列 900 と複数の行 902 と、ブレースホルダー 904 と、ロードされたエントリ 906 とを有

す。

間接セルがフル即ち満杯である場合（最後のフィールド 706 を除いて全ての使用可能な位置内にオフセットを格納済）、間接セルプロセサは次の間接セルが最後のフィールド 606 内に位置されるメモリ内の位置と関連してそのオフセットを格納する。その後、フル即ち満杯の間接セルがメモリ内のその適宜の場所に書込まれる。間接セルのメモリに対する書込は、関連する入力ポート 107 からの入力スイッチ 100 による間接セルブレースホルダーの受取りと一致する。このプロセスはバケット内の最後のセルがメモリ内に格納されるまで継続して行われる。その時に、最後の間接セルがメモリへ書込まれ、且つキバプファ 516 からの関連するエントリ 6

00 が処理するために制御器 106 へ転送される。

しばしばそうであるように、バケットの最後のセルは完了した間接セルをメモリ内にすぐ書込むために必要とされるタイミングと一致するものではない。何故ならば、バケット長は完全に任意的なものであるからである。バケットの最後は全体的な間接セルのファイリングと一致する蓋然性はない。バケットが完了し（全てのセルが入力スイッチによって受取られている）且つ間接セル内の最後のエントリが書込まれる場合には、間接セルはメモリに書込まれるのに自由である。然しながら、その書込は、適切な時間まで遅延され、従って遅延間接セルと呼ばれる所以である。遅延間接セルは、バケットと関連する最後の間接セルである間接セルである。それが遅延される理由は、それは、バケットの残部がメモリに書込まれた後にメモリに書込まれるからである。メモリに対する書込のタイミングは、間接セルに対して予約されているアドレスによって支配される。上述したように、間接セルの作成時にメモリ内の位置が予約される。そのバケットが完了された後に特定のメモリバンクに対して書込むためにその特定の入力ポートに対して使用可能な次のタイムスロットにおいてメモリに対して書込まれる。遅延間接セルのメモリに対する書込のタイミングは、適宜の入力ポー

ト 107 からの遅延間接ブレースホルダーの受取りと一致する。

している。各行は単一のセルスロット上で発生されるべき 1 組の読取要求（メモリバンク当たり 1 個）を表わす。各行は各出力ポート 108 に対する単一のエントリを有している。各セル

スロットにおいて、各出力ポートはメモリ 104 における単一のメモリバンク 105 からの読取を要求することが可能である。読取ポインタ 908 は予約テーブル 508 と関連している。該ポインタは、読取られるべき予約テーブル内の次の行に対してポイントする。該読取ポインタの前の行は後のセルスロットタイムにおいてキュー即ち行列化される要求に対応している。1 実施例においては、該ポインタは各セルスロットタイムにおいて少なくとも 1 個の行を移動する。

ロードされているエントリ 906 は出力スイッチ 102 から受取られた読取要求の結果として実施されるべき読取要求を反映する。ブレースホルダー 904 はいまだに要求されていない使用可能な読取要求を表わす。各セルスロットにおいて、読取制御器 517 は 3 つの機能、即ちテーブル内の最初の使用可能な位置（読取ポインタの後）においてエントリを予約テーブル内にローディングすること、最後の行を出力プロセサ 505 に対する読取要求として出力すること、及び該テーブルをリフレッシュし、最後の行を外へ移動させ、該行をインクリメントし且つテーブルのトップに新たな行を作成すること、を実行する。予約テーブル内の行数は、読取要求を処理する場合の待ち時間とバンク数を掛けた積と同じ

大きさのものでなければならない。1 実施例においては、予約テーブル 508 内に 48 個の行が設けられており、それは 6 個のセルスロットの待ち時間と 8 個のメモリバンクを有するシステムを反映している。

初期化において予約テーブル 508 は行 902 の全てにおいてブレースホルダー 904 を包含している。ブレースホルダー 904 はロードされていない予約テーブル内の位置である。読取要求が読取プロセサによって処理されると、ブレースホルダー 904 のあるものが読取要求に基づいてロードされたエントリ 906 へ変換される。ロードされたエントリ 906 は読取要求アドレスを有している。

次に、図 10 を参照すると、予約テーブルをローディングするプロセスは、出

カポートから読取要求（フルアドレス）を受取ることを包含している（1000）。読取制御器は、この読取要求をデコードして、サーチすべき予約テーブル内の列（それから読み取られるべきメモリバンクに基づいて）を決定する（1002）。読取プロセッサは、予約テーブルの底部から開始して、読取要求が発生した出力ポートに関連する最初のブレースホルダーをサーチする（1004）。読取プロセッサは、その位置に読取要求のフルアドレスを書込むことによってブレースホ

ルダー904をロードしたエントリ906へ変換する（1006）。このプロセスは、読取制御器によって受取られた各読取要求に対して繰り返行われる（1008）。

次に、図11を参照すると、メモリ104は複数のメモリバンク105を有している。各メモリバンクは入力ポート1102と出力ポート1104とを有している。各セルスロットにおいて、各メモリバンクは、入力ポート1102を介して、高々1個の書込及び1個の読取要求を受取る。書込要求は、入力ポート107から受取られたセルと関連している。読取要求は、メモリバンク105から出力スイッチ102へ転送されるべきセルデータに対する要求を反映する。メモリ104から出力スイッチ102へ書込まれたセルに関連するデータ構造を図12に示してある。各セル1200は、出力ポート識別子1202とセルデータ1204とを有している。

1実施例においては、該メモリは、複数のバンクに分割され、その場合にメモリバンクの数は入力ポートと出力ポートとの数に等しい。入力ポートと、出力ポートと、メモリバンクとの間には1対1の関係が存在している。この実施例においては、入力スイッチ100からメモリ104へのセルの転送は、時分割多重態様で行われる。即ち、与えられた入力

ポートからの連続したセルは異なるメモリデスティネーション位置へ指向される。各時間期間（セルスロット）において、入力スイッチは、各入力ポートからの単一のセル（使用可能な場合）をメモリ内へ転送する。次の時間T+1において

00は、マスク1402と、次のホップインデックス（next hop index）ポイン

タ1404と、フルアドレス1406と、オフセット1408と、パケット長1401とを有している。

マスクフィールド1402は、出力スイッチ102へ接続されているどの出力ポートがそのパケットを転送すべきかを表わすために使用される。1実施例においては、該通知は、1つを超える出力ポートへ送ることが可能であり、その結果関連するパケットのブロードキャスト即ち同報通信が行われる。

メモリが各出力ポート108と関連している。次のホップインデックスポインタは該メモリ内の1つの位置をポイントする。該メモリは、パケット転送の特定のタイプと関連するメディアヘッダー情報を格納するために使用される。次のホップアドレス及びメディアヘッダーは出力ポート108に関連して以下に更に詳細に説明する。

フルアドレス1406は、そのパケット内の最初のセルが格納されているメモリ内の開始アドレスを表わす。前に説明したように、オフセット1408はセル又はそのパケットと関連している間接セルを検索するためのリンク情報を与える。

次に、図15を参照すると、出力スイッチは、制御器インターフェース1500と、1個又はそれ以上のメモリ入力1502（1502-0乃至1502-7、各メモリバンクに対して1個）と、1個又は

それ以上の出力1504（1504-0乃至1504-7、各出力ポートに対して1個）と、結果プロセッサ1506と出力プロセッサ1508とを有している。出力スイッチ102は4つの機能、即ち出力結果の受取り、出力結果の処理、メモリからのセルの受取り、出力ポートへのセルの出力、を実行する。

メモリからのセルはメモリ入力1502において受取られ且つ出力プロセッサ1508へ転送される。出力プロセッサ1508はメモリから受取られたセル情報からデスティネーション出力ポートをデコードし且つそのセルデータを適宜の出力

、入力スイッチは、各入力ポートからメモリ内へ再度単一のセルを転送する。同一の入力ポートからの連続するエントリはメモリ104内の異なるメモリバンク105へ書込まれる。

次に、図13を参照すると、制御器106は、制御器メモリ109と、ルートルックアップエンジン110と、入力スイッチインターフェース1300と、出力スイッチインターフェース1302とを有している。制御器106は、入力スイッチインターフェース1300において入力スイッチ100からルートルックアップ要求を受取る。本発明の1実施例においては、複数のルートルックアップエンジン110が制御器106内に設けられており、その各々はルーチング即ち経路付けプロセスを高速化させるためにラウンドロビン態様でルックアップ要求を受取る。1実施例においては、制御器メモリ109はフル即ち満杯の蓄積容量でサービス即ち処理するために32個のルートルックアップエンジン110を必要とする4バンクスタティックランダムアクセス

メモリ（SRAM）である。ルーターを介してのベストマッチのルートを決定する前にパケットから検索されたキーのマッチングは、1996年12月16日付けで出願した発明者がFerguson et al.の「スイッチング装置における高速可変長ベストマッチルックアップ（HIGH SPEED VARIABLE LENGTH BEST MATCH LOOK-UP IN A SWITCHING DEVICE）」という名称の同時係属中の特許出願、出願番号08/767,576号により詳細に記載されており、尚、それを引用によって明示的に本明細書に取込む。

ルートルックアップ要求を処理するルートルックアップエンジンは、ベストマッチ即ち最良の位置のルックアップを実行し且つ出力スイッチ102に対して出力スイッチインターフェース1302を介して通知を出力する。その通知は、パケットをそのデスティネーションへ転送する場合に使用すべき出力ポートを表わす結果を有している。

次に、図14を参照すると、制御器106によって出力スイッチ102へ出力された通知と関連するデータ構造が示されている。この通知用のデータ構造14

1502へ転送する。各セルスロットにおいて、出力スイッチ102はメモリ104内の各バンクから処理するためのセルを受取る。

出力スイッチ102は制御器インターフェース1500を介して制御器106からの通知を受取る。結果プロセッサ1506はその結果（ルート）をデコードし且つどのポート108がそのルートデータを受取るべきかを決定する。その通知内のマスク1402に基づいて、結果プロセッサ1506はその通知をそのようにして示された各夫々の出力ポート108へ転送するために出力ポート1508へ転送する。各セルスロットにおいて、出力プロセッサ1508は（出力1504を介して）ルートを各出力ポート108

へ供給する。

出力プロセッサ1508から出力ポート108へ転送されたデータに関連するデータ構造を図16に示してある。セル1600はヘッダー1602とデータフィールド1604とを有している。ヘッダー1602はメモリバンクソース情報1606とルート情報1608とを有している。メモリバンクソース情報は、どのメモリバンクがデータフィールド1604内にそのセルを供給したかを表わすソース識別子を有している。ルート情報1608は、次のホップインデックスと、パケット長と、フルアドレスと、オフセットとを包含する該通知からのデータを包含している。

次に、図17を参照すると、各出力ポート108は、出力スイッチインターフェース1700と、入力スイッチインターフェース1702と、バッファ1704と、出力要求プロセッサ1706と、ライン出力インターフェース1708と、格納装置（メモリ）1710と、出力バッファ1712と、出力フォーマッター1714とを有している。

出力ポート108はパケットが出力スイッチインターフェース1700において受取られたセル1600によって処理されるべきであることの通知を受取る。出力要求プロセッサ1706はその要求をバッ

ファ1704内に格納し且つ、その後、そのパケットが格納されているメモリ

内の最初のアドレスと関連する入力スイッチ100に対して読取要求を発生する。出力要求プロセッサ1706は出力スイッチ102から受取ったフルアドレスに基づいて最初の読取要求を発生する。その後、間接の読取要求はその要求（セル1600から）又は間接セル（以下に説明するように）において与えられたオフセット情報に基づいて入力スイッチへ転送するために発生される。

セル1600と共に与えられたルート情報から決定されるように、バケット長が5個のセルよりも大きい場合には、出力要求プロセッサは、最初に、そのバケットと関連する最初の間接セルの転送（メモリからの読取）を要求する。このことは、セル1600内に与えられたオフセット及びフルアドレスに基づいて間接セルのアドレスを計算することによって行われる。間接セル要求が発生された後に、出力要求プロセッサはセル1600内に与えられたオフセットとフルアドレスに基づいてバケット内の残りのセルに対する読取要求を発生する。出力スイッチ102から間接セルを受取ると、出力要求プロセッサは間接セル内に包含されているオフセット情報に基づいてそのバケット内の残りのセルに対する読取要求を継

続して発生する。

その後の間接セルは同様の態様で検索される。即ち、次の間接セルの読取を行う時に、次の間接セルのアドレスが前の間接セル内に格納されている最後のオフセットに基づいて計算される。間接セルを検索するタイミングは、出力ストリーム内における遅延が必要とされないように実施される。各その後の間接セルは先の間接セルの終了前に検索される。このように、出力ストリームが初期化されると、データのバッファ動作が必要とされることはなく、且つ検索プロセッサに関連する待ち時間に起因する中断を経験することはない。

個々のメモリバンクに対する出力要求は厳格に順番通りに処理される。即ち、出力ポートはメモリバンクに対して発行された各要求をトラッキング即ち追跡することが可能であり且つ同一のメモリバンクに対する一連の要求に回答して受取られたデータはそれらが発行されたシーケンス又はパターンに従って厳格に送給される。出力要求プロセッサは各メモリバンクに対して発生された要求を追跡する。出力バッファ712は複数個のキューを有しており、即ち各メモリバンクに対

して入力ポートはそのバケットを固定長セルへ分割し且つ該セルを入力スイッチへ転送する（1802）。入力スイッチは、バケット内の最初のセルからキー情報を除去し且つそれを一時的にキーバッファ内に格納する（1804）。その後、本入力スイッチは、該セルを時分割多重化態様でメモリバンクへルーティング即ち経路付けさせる（1806）。本入力スイッチは最初のセルがメモリ内のどこに書込まれるかに関連する最初のアドレスを格納し且つ次のセルが書込まれる次の隣接したメモリバンクに

対してメモリ内のオフセットに関連する各付加的なセルに対するオフセットを計算する（1808）。本入力スイッチはバケット長が5個のセルを超えるものである場合にそのバケットに対するリンク情報を格納するための間接セルを作成する（1810）。そのセル数が間接セル内における使用可能なオフセットの数を超える場合には、古い間接セルがメモリ内に格納され且つ新たな間接セルが作成され且つ入力スイッチにおいて受取られた各新たなセルに対して計算されたオフセットに基づいてロードされる。

バケット（及び、存在する場合には、その間接セル）がメモリ内に格納されると、キーと、第一セルのフルアドレスと、オフセット情報とがルックアップ要求として制御器へ転送される（1814）。該制御器はベストマッチのルックアップを実行し且つそのルックアップの結果を発生する。その結果は、デスティネーションポート（出力ポート）とアドレスと、オフセット情報と、次のホップインデックスとを有している（1816）。その結果を包含する通知は適宜の出力ポートへ転送するために出力スイッチへ転送される（1818）。

通知を受取ると、出力ポートはそのバケットと関連するデータに対して入力スイッチに対し一度に1個のセルの読取要求を発生する（1820）。入力

スイッチは、時分割多重態様でその読取要求を発行し、セルスロット当たり各メモリバンクに対して単一の要求を発生する（1822）。メモリバンクが入力スイッチからの要求を受取ると、その要求及びセルデータと関連している出力ポートが出力スイッチに対して書込まれる（1824）。再度、各セルスロットにお

して2個設けられており（1個の要求キューと1個の応答キュー）、該キューは、それらが割当てられている特定のストリームに従っ

てメモリから受取られたセルの順番付けを行うために使用される。要求キューはストリーム番号と読取アドレスとを包含している。メモリに対して要求が発行されると、その要求キューからエントリが除去され且つストリーム番号部分が応答キュー内に配置される。応答が受取られると、応答キューのヘッドにおけるエントリが除去され且つその応答が応答キューから検索されたストリーム番号によって表わされるストリーム番号へ送られる。

出力ポート108においてセルが受取られると（読取要求に回答して）、それらは出力バッファ1712内に格納される。与えられたバケットに対して、出力ポートはストリーム化した出力を与えることを必要とされるセルの数を格納する。本発明の1実施例においては、出力ポートからの出力（ストリームデータ）を開始する前に、12個のセルが格納される。出力バッファ1712内に格納するセルの数の選択は読取プロセスにおける待ち時間に基づくものである（出力ポートからの読取要求とその読取要求と関連しているセルの出力ポートへの到着との間のクロックサイクル数）。

出力フォーマッタ1714は出力バッファ1712からセルを受取り且つそのデータをメモリ1710内に格納されているメディアヘッダー情報と結合さ

せる。出力スイッチ102から受取られた各要求（通知）は次のホップインデックス（next hop index）を有している。この次のホップインデックスは、与えられたタイプの送信（バケットのデスティネーションから派生される）と関連するメディアヘッダー情報のメモリ1710内の開始アドレスを表わす。出力フォーマッタ1714はメモリから戻されたセルデータを適宜のメディアヘッダーと結合してライン出力インターフェース1708を介してルーター20から転送するための適切なバケットを発生する。

次に、図18を参照すると、スイッチを介してバケットをルーティング即ち経路付けする方法において、バケットが入力ポートにおいて受取られる（1800）

いて、出力スイッチは単一のセルを出力ポートの各々へ転送する。受取ると、出力ポートはそのセルデータをメディアヘッダー情報と結合し且つそのデータをデスティネーションへストリームさせる（1826）。

変形実施例

本発明の1実施例においては、出力ポートと入力ポートと、メモリバンクとが単一の装置内に包含されている。このメモリを包含する多機能ポートのアーキテクチャを図19に示してある。特に、多機能ポート1900はメモリバンク105と、ライン入力インターフェース300と、データハンドラー302と、バッファ1704と、出力要求プロセッサ1706と、ライン出力インターフェース1708と、格納装置1710と、FIFO1712と、出力フォーマッタ1714と、入力スイッチインターフェース1902と、出力スイッチインターフェース1904とを有している。

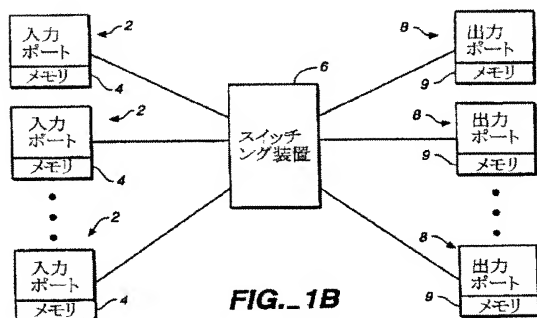
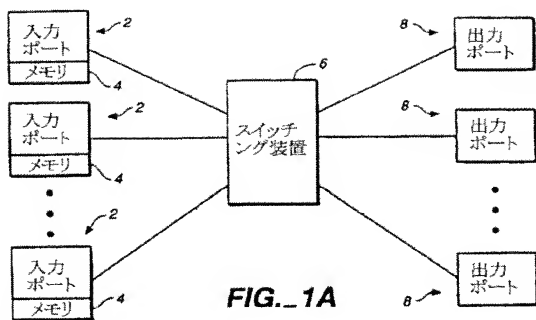
この多機能ポートは図20に示したように入力スイッチ、出力スイッチ及び制御器に関連して使用される。入力ポート、出力ポート及びメモリバンクの種々の部分的な構成要素は上述した個別的な機能を実行する。然しながら、これらの装置を単一のユニットへ結合させることはコンポーネント間のインターフェースを簡単化させる。特に、多機能ポートと入力スイッチとの間での転送に対する改定したフォーマットを図21に示してある。

多機能ポート1900から入力スイッチへ転送されたセル2100はセルヘッダー2102とセルデータ2104とを包含している。セルヘッダー2102は、図4を参照して上述したフィールドと機能が同様の、タイプフィールド406と、ストリームフィールド408と、バケットヘッダーフィールド410とを包含している。更に、セルヘッダー2102は出力ポート識別子2106及びアドレス2108の形態での読取要求を包含している。出力ポート識別子2106はその読取要求を発生する出力ポートを識別する。アドレス2108は読取られるべきメモリ104内のアドレスを表わす。

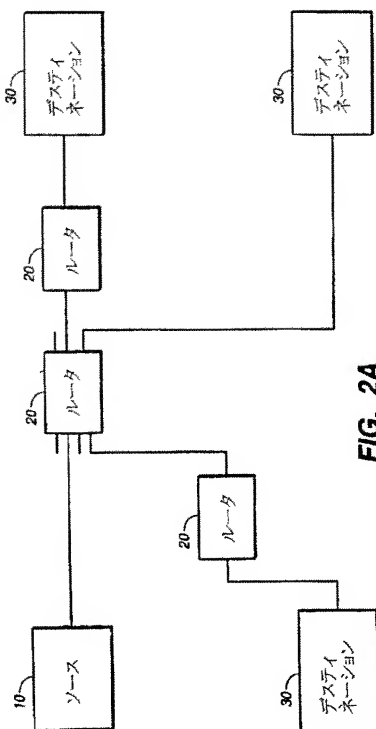
本発明を特定の実施例について説明したが、それらは本発明の例示的なものであって制限的なものとして解釈されるべきものではない。その他の実施例

は以下の請求の範囲内のものである。

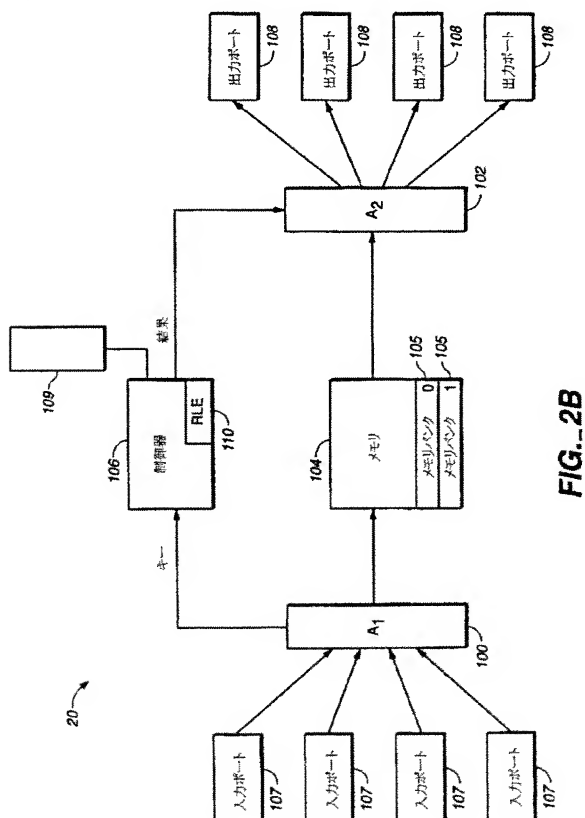
【図 1】



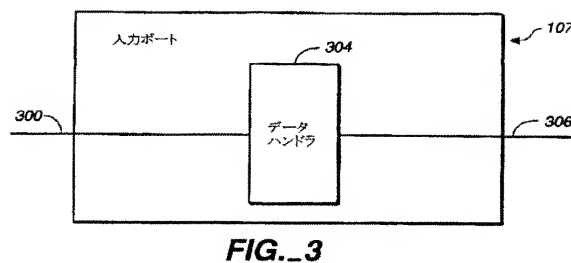
【図 2 A】



【図 2】



【図 3】



【図 4】

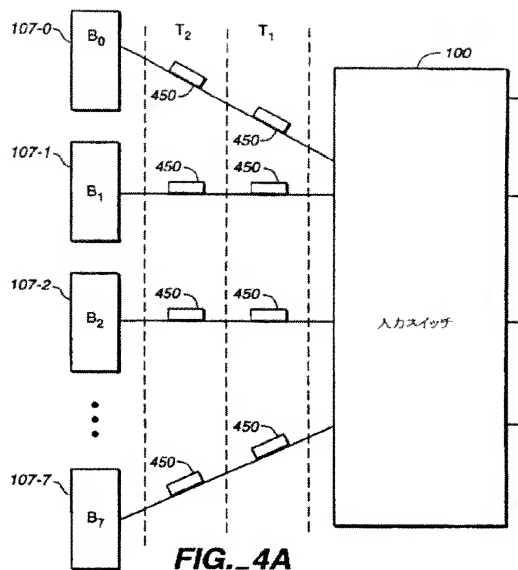


FIG. 4A

【図 4】

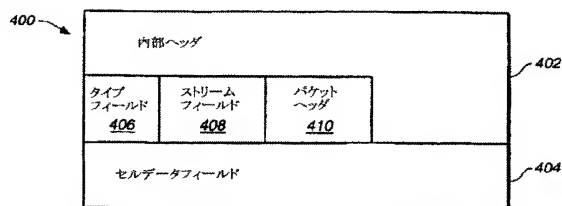


FIG. 4B

【図 5】

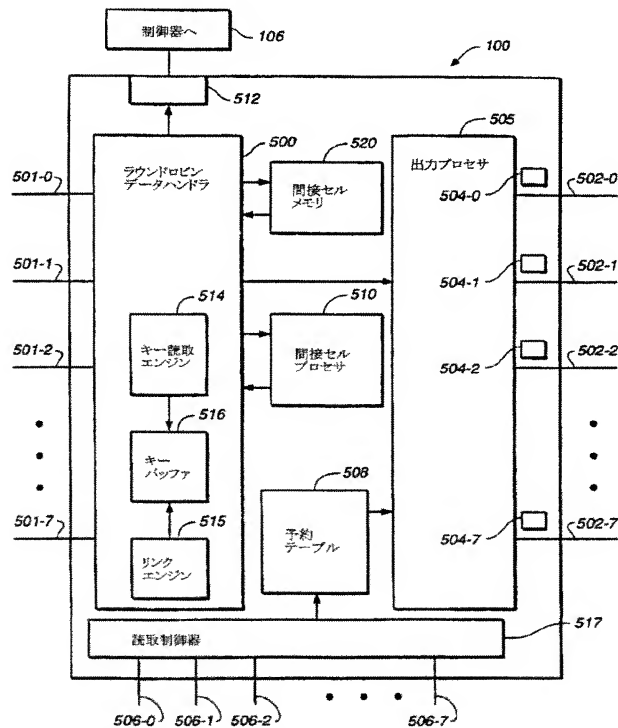


FIG. 5A

【図 6】

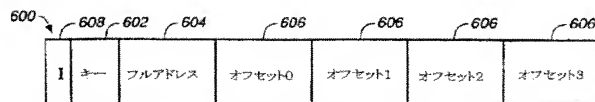


FIG. 6

【図 7】

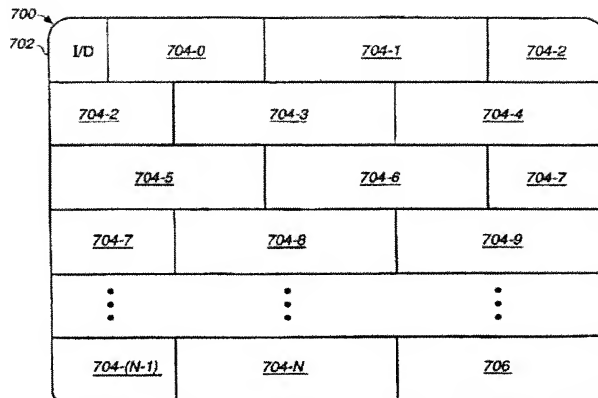


FIG. 7

【図 5】

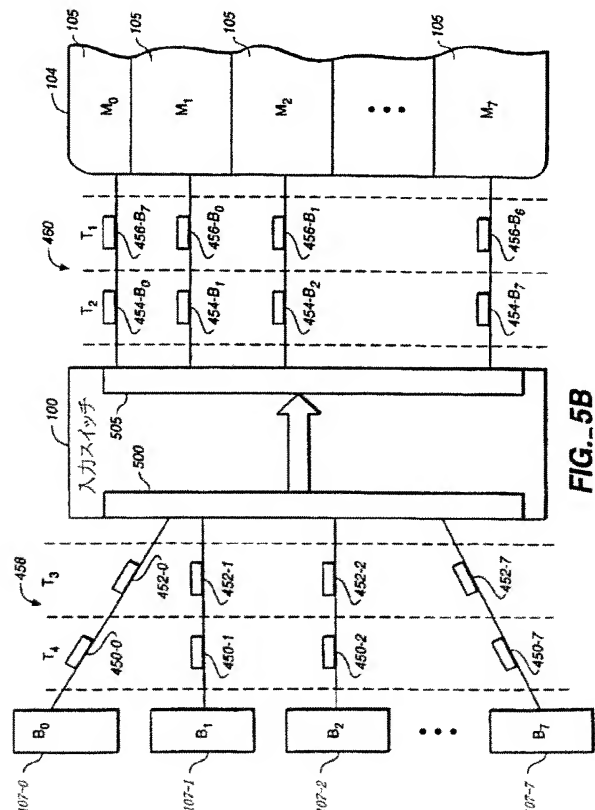


FIG. 5B

【図 8】

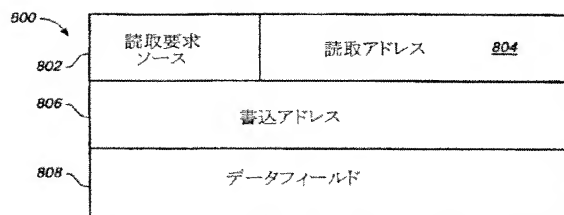


FIG._8

【図 9】

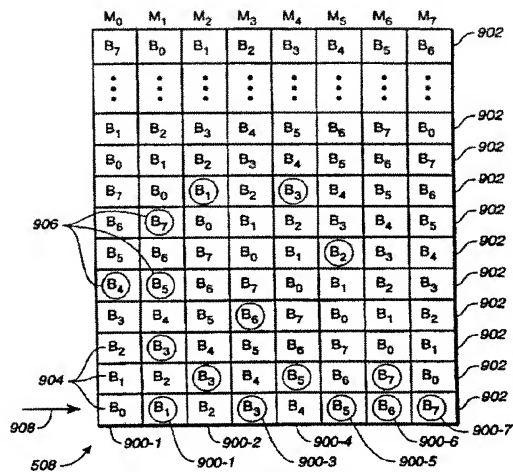


FIG._9

【図 10】

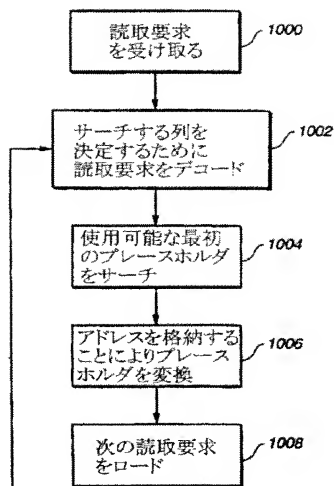


FIG._10

【図 11】

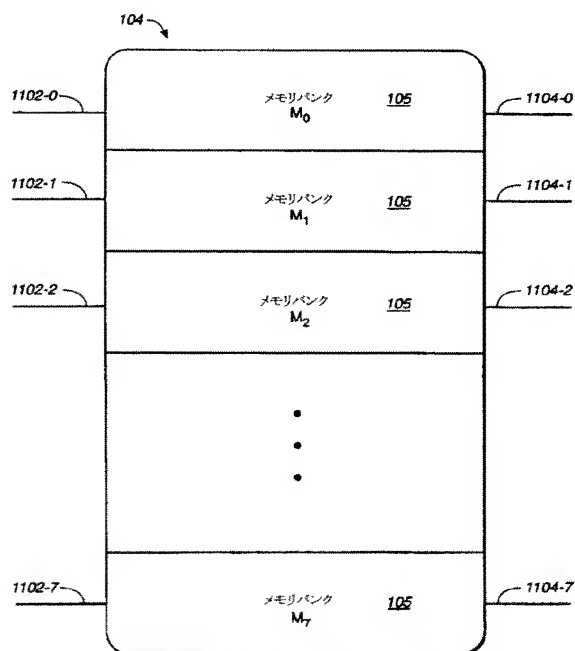


FIG._11A

【図 11】

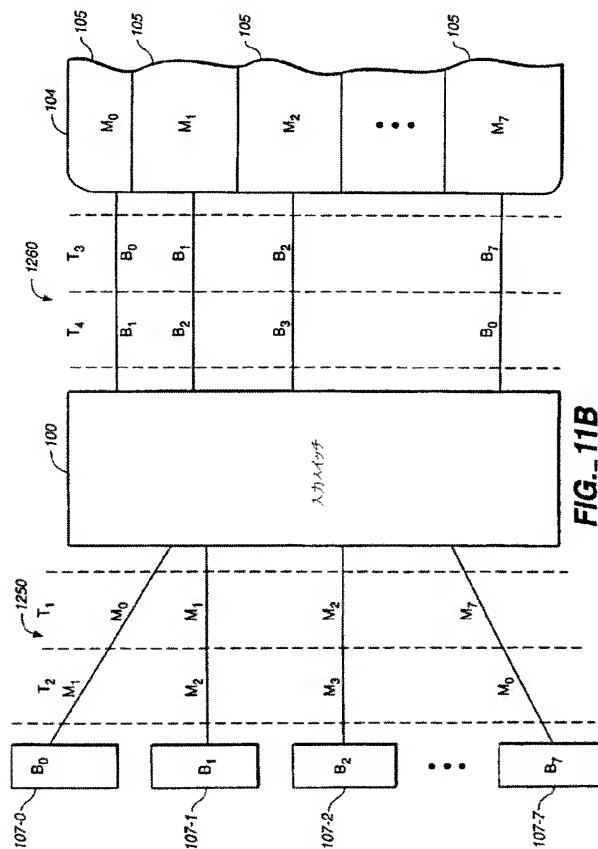


FIG._11B

【図 12】

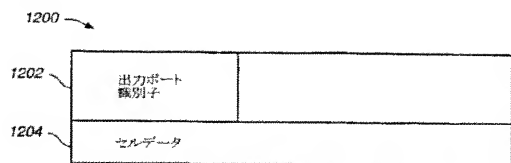


FIG._12

【図 13】

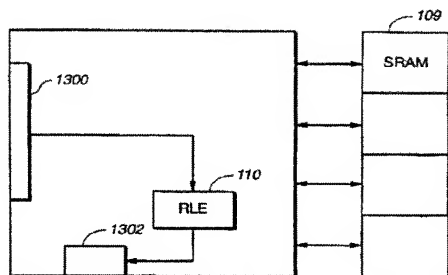


FIG._13

【図 14】

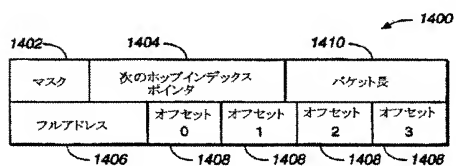


FIG._14

【図 16】

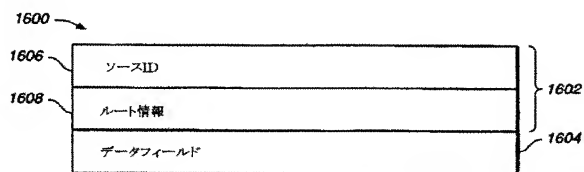


FIG._16

【図 15】

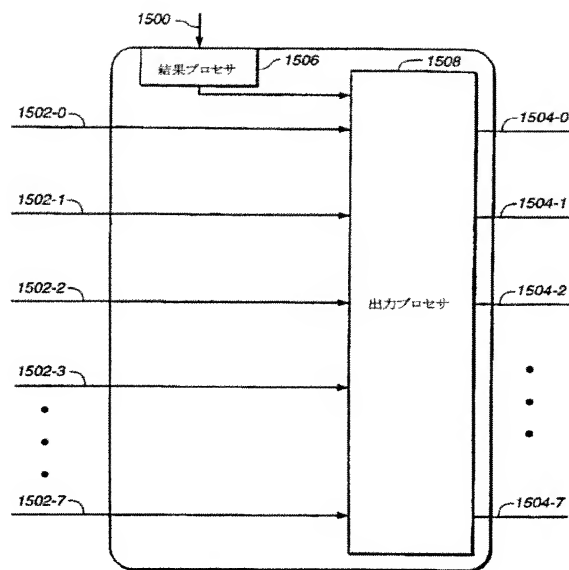


FIG._15

【図 17】

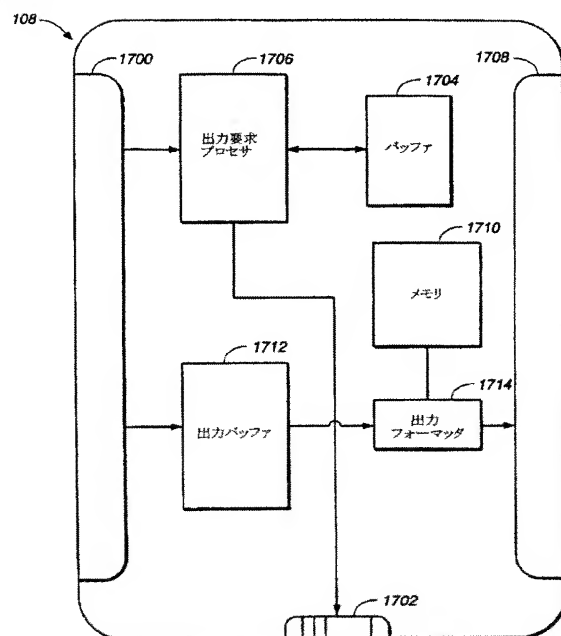


FIG._17

【図18】

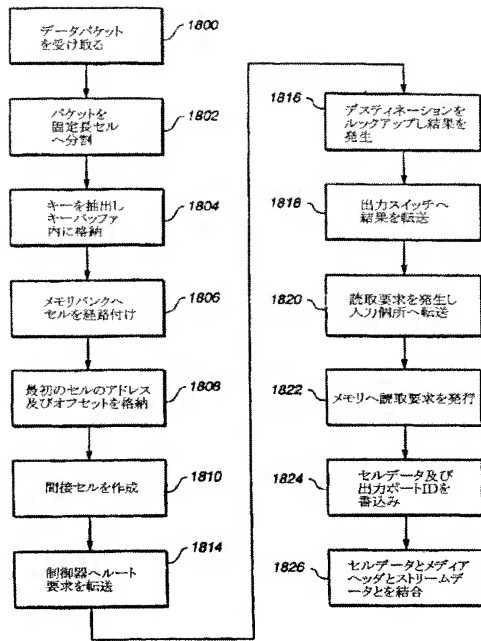


FIG. 18

【図19】

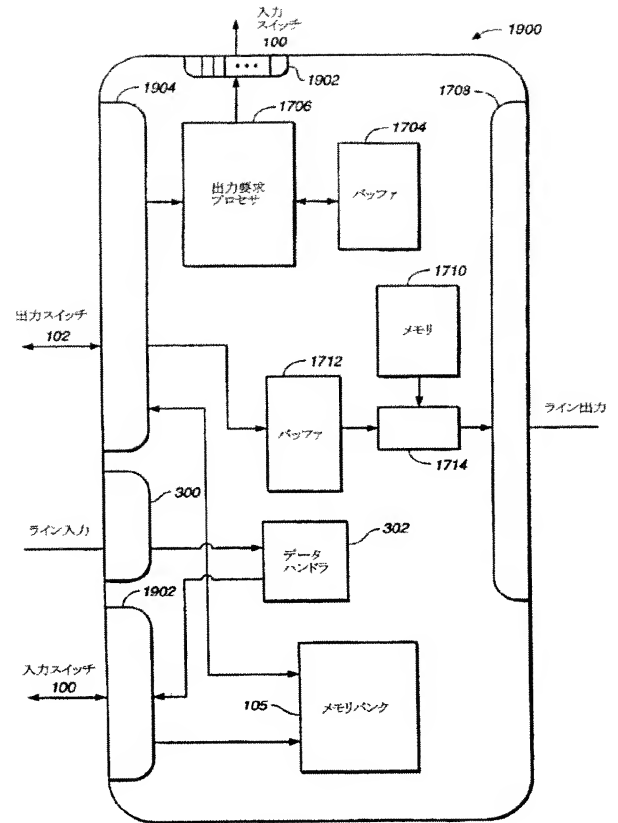


FIG. 19

【図20】

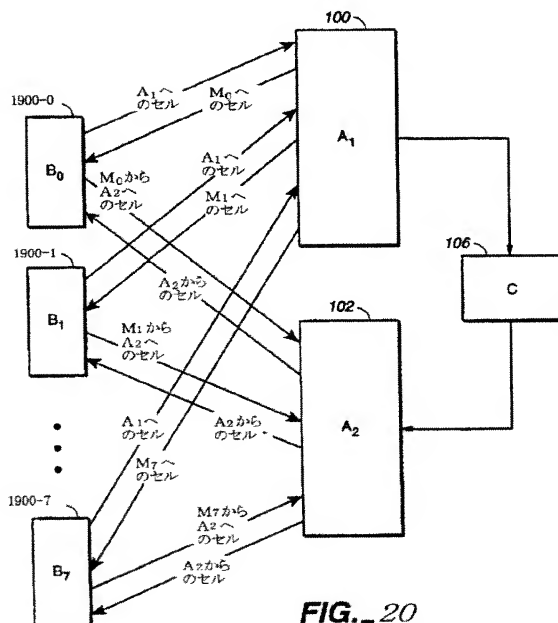


FIG. 20

【図21】

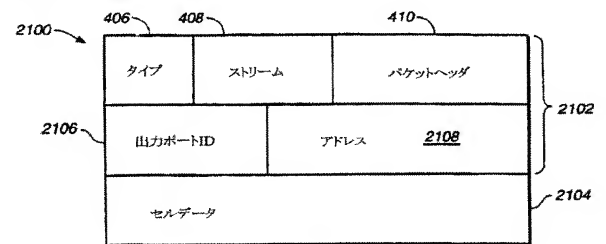


FIG. 21

フロントページの続き

(31)優先権主張番号 08/901,061

(32)優先日 平成9年7月24日(1997. 7. 24)

(33)優先権主張国 米国(US)

(81)指定国 EP(AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), CA, JP

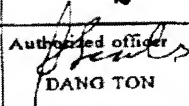
(72)発明者 ファーガソン, デニス シイ.
アメリカ合衆国, カリフォルニア 94043,
マウンテン ビュー, オーチャード グレ
ン コート 203

(72)発明者 レインクレス, ビョルン オー.
アメリカ合衆国, カリフォルニア 94303,
パロ アルト, グリアー ロード 2731

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/23285

A. CLASSIFICATION OF SUBJECT MATTER IPC(6) : H04L 12/56 US CL : Please See Extra Sheet. According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 370/389,401,402,407,412,413,414,415,416,417,428,429; 395/280,844,873,917,909; 364/238,240,221 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched NONE Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) NONE		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,491,694 A (OLIVER ET AL.) 13 February 1996, see columns 1-6 and figures 1 and 3.	1,4-12,14,17-25
Y		2-3,13,15-16
Y	US 5,448,702 A (GARCIA ET AL) 05 September 1995, see column 6 lines 20-68.	2-3,13,15-16
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: *A* document defining the general state of the art which is not considered to be of particular relevance *B* earlier document published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art *Z* document member of the same patent family		
Date of the actual completion of the international search 09 APRIL 1998		Date of mailing of the international search report 28 JUL 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer  DANG TON Telephone No. (703) 305-4739

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/23285

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

370/389,401,402,407,412,413,414,415,416,417,428,429; 395/280,844,873,917,909; 364/238,240,221